

Focal-plane wavefront sensing with the vector Apodizing Phase Plate

S.P. Bos¹, D.S. Doelman¹, J. Lozi², O. Guyon^{2,3,4,5}, C.U. Keller¹, K.L. Miller¹,
N. Jovanovic⁶, F. Martinache⁷, F. Snik¹

¹ Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands
e-mail: stevenbos@strw.leidenuniv.nl

² National Astronomical Observatory of Japan, Subaru Telescope, National Institute of Natural Sciences, Hilo, HI 96720, USA

³ Steward Observatory, University of Arizona, 933 N. Cherry Ave, Tucson, AZ 85721, USA

⁴ College of Optical Sciences, University of Arizona, 1630 E. University Blvd., Tucson, AZ 85721, USA

⁵ Astrobiology Center, National Institutes of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo, Japan

⁶ Department of Astronomy, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

⁷ Observatoire de la Cote d’Azur, Boulevard de l’Observatoire, Nice, 06304, France

Received June 12, 2019; accepted September 16, 2019

ABSTRACT

Context. One of the key limitations of the direct imaging of exoplanets at small angular separations are quasi-static speckles that originate from evolving non-common path aberrations (NCPA) in the optical train downstream of the instrument’s main wavefront sensor split-off.

Aims. In this article we show that the vector-Apodizing Phase Plate (vAPP) coronagraph can be designed such that the coronagraphic point spread functions (PSFs) can act as a wavefront sensor to measure and correct the (quasi-)static aberrations, without dedicated wavefront sensing holograms nor modulation by the deformable mirror. The absolute wavefront retrieval is performed with a non-linear algorithm.

Methods. The focal-plane wavefront sensing (FPWFS) performance of the vAPP and the algorithm are evaluated with numerical simulations, to test various photon and read noise levels, the sensitivity to the 100 lowest Zernike modes and the maximum wavefront error (WFE) that can be accurately estimated in one iteration. We apply these methods to the vAPP within SCExAO, first with the internal source and subsequently on-sky.

Results. In idealised simulations we show that for 10^7 photons the root-mean-square (RMS) WFE can be reduced to $\sim \lambda/1000$, which is 1 nm RMS in the context of the SCExAO system. We find that the maximum WFE that can be corrected in one iteration is $\sim \lambda/8$ RMS or ~ 200 nm RMS (SCExAO). Furthermore, we demonstrate the SCExAO vAPP capabilities by measuring and controlling the lowest 30 Zernike modes with the internal source and on-sky. On-sky, we report a raw contrast improvement of a factor ~ 2 between 2 and $4 \lambda/D$ after 5 iterations of closed-loop correction. When artificially introducing 150 nm RMS WFE, the algorithm corrects it within 5 iterations of closed-loop operation.

Conclusions. FPWFS with the vAPP’s coronagraphic PSFs is a powerful technique since it integrates coronagraphy and wavefront sensing, eliminating the need for additional probes and thus resulting in a 100% science duty cycle and maximum throughput for the target.

Key words. Instrumentation: adaptive optics—Instrumentation: high angular resolution

1. Introduction

The exploration of circumstellar environments at small angular separations by means of direct imaging is crucial for the detection and characterization of exoplanets. The challenges that need to be overcome are that of high contrast and small angular separation: Earth around the Sun at 10 pc will have an angular separation of ~ 100 mas and a contrast in the visible ($\sim 0.3\text{--}1 \mu\text{m}$) of $\sim 10^{-10}$ (Traub & Oppenheimer 2010).

The current generation of ground-based high-contrast imaging instruments SPHERE (Beuzit et al. 2019), GPI (Macintosh et al. 2014), SCExAO (Jovanovic et al. 2015) and the upcoming MagAO-X (Males et al. 2018; Close et al. 2018) are

pushing towards contrasts of $\sim 10^{-6}$ at angular separations of 200 mas after post-processing in the near-infrared (0.95–2.3 μm ; Vigan et al. 2015). These instruments are equipped with extreme adaptive optics systems to flatten the wavefront after the turbulent atmosphere, coronagraphs to suppress the star light and contrast-enhancing post-processing techniques that employ some form of diversity such as angular differential imaging (Marois et al. 2006), reference star differential imaging (Ruane et al. 2019), spectral differential imaging (Sparks & Ford 2002) and polarimetric differential imaging (Snik & Keller 2013). The latter two techniques can also be used as a characterization diagnostic. Medium- and high-resolution integral-field spectroscopy

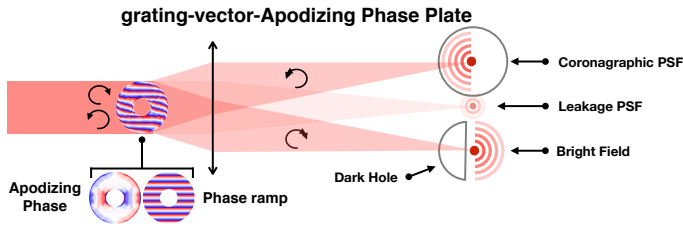


Fig. 1. Working principle of the grating-vector-Apodizing Phase Plate. The vAPP is a half-wave retarder pupil-plane optic with a spatially varying fast-axis orientation. The varying fast-axis orientation induces the phase through the geometric phase on the circular polarization states. These polarization states receive the opposite phase and therefore have flipped coronagraphic point spread functions (PSFs). The PSFs are spatially separated by adding a phase ramp to the design. Any offsets from half-wave retardance within the optic reduces the efficiency with which the light will be transferred to the coronagraphic PSFs and results in a leakage PSF.

can be used to detect atomic and molecular lines from a planet's atmosphere (e.g. Haffert et al. 2019 and Hoeijmakers et al. 2018) while polarimetry can be used to detect cloud structures (Stam et al. 2004; De Kok et al. 2011; van Holstein et al. 2017).

The coronagraph relevant for the present work is the vector-Apodizing Phase Plate (vAPP; Snik et al. 2012; Otten et al. 2017), which manipulates the phase in the pupil-plane such that in selected regions in the focal-plane the starlight is cancelled; these areas are referred to as the dark holes. The phase is induced through the achromatic geometric phase (Pancharatnam 1956; Berry 1987) on the circular polarization states by a half-wave liquid-crystal layer with a varying fast-axis orientation. The two circular polarization states both receive equal, but opposite phases, resulting in two coronagraphic PSFs with opposite dark holes as shown in Figure 1. The geometric phase is inherently achromatic as it depends on geometric effects, but the efficiency with which the light acquires the phase is determined by the retardance of the liquid-crystal layer. Retardance offsets from half-wave result in leakage; light that has not acquired the desired phase will form a non-coronagraphic PSF based on the aperture geometry. Generally, vAPP coronagraphs are designed to have minimal leakage over a broad wavelength range. High leakage will affect coronagraphic performance as light from the leaked PSF will contaminate the dark hole. In the simplest and most common implementation the two coronagraphic PSFs are spatially separated with a polarization-sensitive grating (Oh & Escuti 2008) that is integrated in the phase design. These coronagraphs are known as grating-vAPPs (Otten et al. 2014) and are mainly used for operation with narrowband filters or integral-field spectrographs to prevent smearing by the grating. In this article the grating-vAPP will be referred to as a vAPP. The vAPP has been put on-sky with several instruments: CHARIS/SCExAO (Doelman et al. 2017), MagAO/Clio2 (Otten et al. 2017), LMIRCAM/LBT (Doelman et al. 2017) and LEXI (Haffert et al. 2018). Furthermore, vAPPs have been designed for the following instruments: HiCIBaS (Côté et al. 2018), MagAO-X (Miller et al. 2018), ERIS (Boehle et al. 2018), METIS (Kenworthy et al. 2018) and MICADO (Davies et al. 2018).

One of the key limitations of current high contrast imaging instruments that limit them to contrasts above $\sim 10^{-6}$ within 300 mas are quasi-static speckles that originate from slowly-evolving instrumental aberrations caused by changing

temperature, humidity and gravity vector during observations (Martinez et al. 2012; Martinez et al. 2013; N'Diaye et al. 2016; Vigan et al. 2018). When these aberrations occur in the optical train downstream of the main wavefront sensor split-off, they will not be sensed and therefore can not be corrected. Additional focal-plane wavefront sensing (FPWFS) with the science detector is a highly desirable solution to these non-common path aberrations (NCPA). Besides eliminating the NCPAs, a FPWS can also address chromatic errors between the main sensing and science channels. Another advantage of FPWFS is that it has been shown by Guyon (2005) to be able to reach high sensitivities for all spatial frequencies, only being bested in sensitivity by the Zernike wavefront sensor (N'Diaye et al. 2013; Doelman et al. 2019). A notable FPWFS is the Self-Coherent Camera (SCC; Baudoz et al. 2005; Galicher et al. 2008; Mazoyer et al. 2013). This is a WFS that is combined with a coronagraph, which uses focal-plane optics to block starlight. It operates by placing a small hole in the pupil-plane Lyot stop outside of the geometric pupil where the scattered starlight is located. This hole creates a reference beam that will generate high spatial frequency fringes in the focal-plane image, which can be used to determine the full electric field. The SCC has a 100% science duty cycle, but requires a high focal-plane sampling to resolve the fringes, and optics of a sufficient size to accommodate the reference beam. FPWS has also been conducted by using vAPPs. Previous work focused on adding additional holograms in the focal-plane that encode wavefront information (Wilby et al. 2017) or directly probe the electric field (Por & Keller 2016); these will not be considered here.

An overview of FPWFS techniques can be found in Jovanovic et al. (2018). Three FPWFS and control methods particularly relevant are:

- The COronagraphic Focal-plane waveFront Estimation for Exoplanet detection (COFFEE; Sauvage et al. 2012; Paul et al. 2013a) wavefront sensor is an extension of classical phase diversity (Gonsalves 1982; Paxman et al. 1992) to coronagraphic imaging. Aberrations in a physical model of the coronagraphic system are fitted to two focal-plane images, one of which has a known phase diversity (e.g. defocus). The method has been demonstrated in the lab (Paul et al. 2013b) and on the SPHERE system using the internal calibration source (Paul et al. 2014). Recent extensions enable COFFEE to measure phase in long-exposure images affected by residual turbulence (Herscovici-Schiller et al. 2017), and measure both phase and amplitude (Herscovici-Schiller et al. 2018).
- An interferometric approach to FPWFS is the the Asymmetric Pupil Fourier Wavefront Sensor (APF-WFS; Martinache 2013). The APF-WFS assumes the small aberration regime to use a Fourier analysis of focal-plane images to determine pupil-plane phase aberrations. The image is formed by an asymmetric pupil, which enables the full phase determination. The theory behind this technique will be more extensively discussed in section 2. The wavefront sensor has been demonstrated on-sky in closed-loop operation controlling the lowest-order Zernike modes (Martinache et al. 2016), also in the context of controlling the "island effect" (N'Diaye et al. 2018).
- The third method is linear dark-field control (LDFC; Guyon et al. 2017) and more specifically the spatial LDFC variant (Miller et al. 2017). The idea behind spatial LDFC is to measure and control small aberrations that pollute the dark hole

(created by the vAPP or other techniques) by measuring the response of the bright field (see Figure 1) relative to a reference state. This response is approximately linear for small aberrations. In Miller et al. (2017) it was shown to work for one-sided dark holes and modes that have a response in the bright field and the dark hole. But, in this case, there also is a spatial null-space consisting of the modes that do not have a response in the bright field but do pollute the dark hole. The authors overcame this problem by using the vAPP (Miller et al. 2018), where the bright field of one coronagraphic PSF covers the dark hole of the other. This has been shown to work in the laboratory (Miller 2018), but still has a significant null space for most vAPP designs - it is insensitive to the even pupil phase modes - as will be shown and addressed in section 2.

COFFEE and the APF-WFS both suffer from low duty cycles as the science observations have to be stopped for the phase diversity probes or moving the asymmetric mask in and out the beam. LDFC, on the other hand, has a 100% duty cycle, but has currently only been considered for vAPPs with a significant null space - the even pupil phase modes - only works in the small aberration regime and does not perform an absolute phase measurement. In section 2 we combine the APF-WFS and LDFC with vAPPs, eliminating the null space and improving the duty cycle to 100%. In section 3 we present a non-linear algorithm similar to COFFEE that can perform absolute phase retrieval. In section 4 we explore the theoretical FPWFS performance of the vAPP and the non-linear algorithm with simulations for the vAPP installed at SCExAO. In section 5 we demonstrate the method first with the internal source and subsequently on sky. In section 6 we discuss the results and present the conclusions.

2. Theory

2.1. Phase retrieval

Phase retrieval techniques in astronomy are concerned with sensing pupil-plane phase aberrations by analysing focal-plane images. These techniques require a unique response in the focal-plane intensity for every phase mode and sign of its modal coefficient (the amount of wavefront error in the specific phase mode). For symmetric pupils such a unique response does not exist for the sign of the modal coefficients of even phase modes. For example, it is easy to determine how an optic needs to be moved to correct for tip/tilt (odd Zernike mode). But when the Point Spread Function (PSF) is defocussed (even Zernike mode), it is not immediately clear in which way the optic needs to be moved to bring the PSF back into focus. We will demonstrate the origin of this well-known sign ambiguity (Gonsalves 1982; Paxman et al. 1992) in this section.

The electric field in the pupil-plane consists of an amplitude component $A(r)$ and a phase component $\theta(r)$:

$$\begin{aligned} E_{\text{pup}}(r) &= A(r)e^{i\theta(r)} \\ &= A(r)\cos[\theta(r)] + iA\sin[\theta(r)]. \end{aligned} \quad (1) \quad (2)$$

Here, the electric field E_{pup} and its constituents, e.g. A and θ , are 2D entities where the position vector r , defined from the center of the pupil, is omitted from here on. The focal-plane electric field $E_{\text{foc}}(x)$ is formed by propagating E_{pup} using the Fraunhofer propagation operator $C\{\cdot\} \propto \frac{1}{i} \mathcal{F}\{\cdot\}$ (Goodman 2005). We use the Fraunhofer propagator, instead of just the Fourier transform,

because it is the physically correct propagator.

$$E_{\text{foc}}(x) = C\{E_{\text{pup}}\} \quad (3)$$

$$= C\{A\cos(\theta)\} + C\{iA\sin(\theta)\} \quad (4)$$

$$= a(x) + ib(x), \quad (5)$$

here $a(x)$ and $b(x)$ are the real and imaginary components of $E_{\text{foc}}(x)$, respectively, and generally contain a mixture of $C\{A\cos(\theta)\}$ and $C\{iA\sin(\theta)\}$. The focal-plane coordinates are denoted by x and are omitted from here on as well. The focal-plane intensity I_{foc} or PSF is subsequently given by:

$$I_{\text{foc}} = |E_{\text{foc}}|^2 \quad (6)$$

$$= |a|^2 + |b|^2 \quad (7)$$

Before we continue with an example of even phase aberrations, recall:

- That a function $f(r)$ can be decomposed into even and odd functions:

$$f(r) = f_{\text{even}}(r) + f_{\text{odd}}(r), \quad (8)$$

$$f_{\text{even}}(r) = \frac{f(r) + f(-r)}{2}, \quad (9)$$

$$f_{\text{odd}}(r) = \frac{f(r) - f(-r)}{2}. \quad (10)$$

An example of such a symmetry decomposition of phase and amplitude in the context of the APP coronagraph is shown in Figure 4.

- The multiplication and composition properties of even and odd functions:

$$f_{\text{even}}(r) \cdot g_{\text{odd}}(r) = h_{\text{odd}}(r), \quad (11)$$

$$f_{\text{odd}}(r) \cdot g_{\text{odd}}(r) = h_{\text{even}}(r), \quad (12)$$

$$f_{\text{even}}(r) \cdot g_{\text{even}}(r) = h_{\text{even}}(r), \quad (13)$$

$$f_{\text{even}}[g_{\text{odd}}(r)] = h_{\text{even}}(r), \quad (14)$$

$$f_{\text{odd}}[g_{\text{odd}}(r)] = h_{\text{odd}}(r), \quad (15)$$

$$f_{\text{odd}}[g_{\text{even}}(r)] = h_{\text{even}}(r), \quad (16)$$

$$f_{\text{even}}[g_{\text{even}}(r)] = h_{\text{even}}(r). \quad (17)$$

- The symmetry properties of Fraunhofer propagation, which are shown in Table 1.
- The Hermitian properties of the Fraunhofer propagation. These say that a conjugated pupil-plane electric field $E'_{\text{pup}} = E_{\text{pup}}^*$ (i.e. a phase sign flip; $*$ denotes the conjugation) will result in a flipped and conjugated focal-plane electric field $E'_{\text{foc}} = C\{E'_{\text{pup}}\}$:

$$E'_{\text{foc}}(r) = -E_{\text{foc}}(-r)^*. \quad (18)$$

The reason that the symmetry decomposition, combined with the decomposition of E_{pup} in its real and imaginary components (Equation 2), is important, is that the Fraunhofer propagation maps combinations of these decompositions into either real or imaginary components of E_{foc} , as shown in Table 1. It will be shown below that these properties of the Fraunhofer propagation determine what kind of symmetry has to be introduced in the pupil-plane to determine the sign of even phase aberrations.

Let's assume that A is even, which is true for most instrument pupils. If an even phase aberration is added, e.g. astigmatism, the terms $A\cos(\theta)$ and $iA\sin(\theta)$ will both be even. Note that

Table 1. Fraunhofer propagation symmetry properties (Goodman 2005), these are the Fourier properties multiplied with a factor $-i$ ($E_{\text{foc}} = C\{E_{\text{pup}}\} \propto \frac{1}{i}\mathcal{F}\{E_{\text{pup}}\}$).

| Pupil-plane electric field $E_{\text{pup}} = A \cos(\theta) + iA \sin(\theta)$ | | | Focal-plane electric field $E_{\text{foc}} = a + ib$ | |
|---|---------------|---|---|---------------|
| Term | Term symmetry | A, θ symmetry | Term | Term symmetry |
| $A \cos(\theta)$ | Even | $(A_{\text{even}}, \theta_{\text{even}}), (A_{\text{even}}, \theta_{\text{odd}})$ | ib | Even |
| $A \cos(\theta)$ | Odd | $(A_{\text{odd}}, \theta_{\text{even}}), (A_{\text{odd}}, \theta_{\text{odd}})$ | a | Odd |
| $iA \sin(\theta)$ | Even | $(A_{\text{even}}, \theta_{\text{even}}), (A_{\text{odd}}, \theta_{\text{odd}})$ | a | Even |
| $iA \sin(\theta)$ | Odd | $(A_{\text{even}}, \theta_{\text{odd}}), (A_{\text{odd}}, \theta_{\text{even}})$ | ib | Odd |

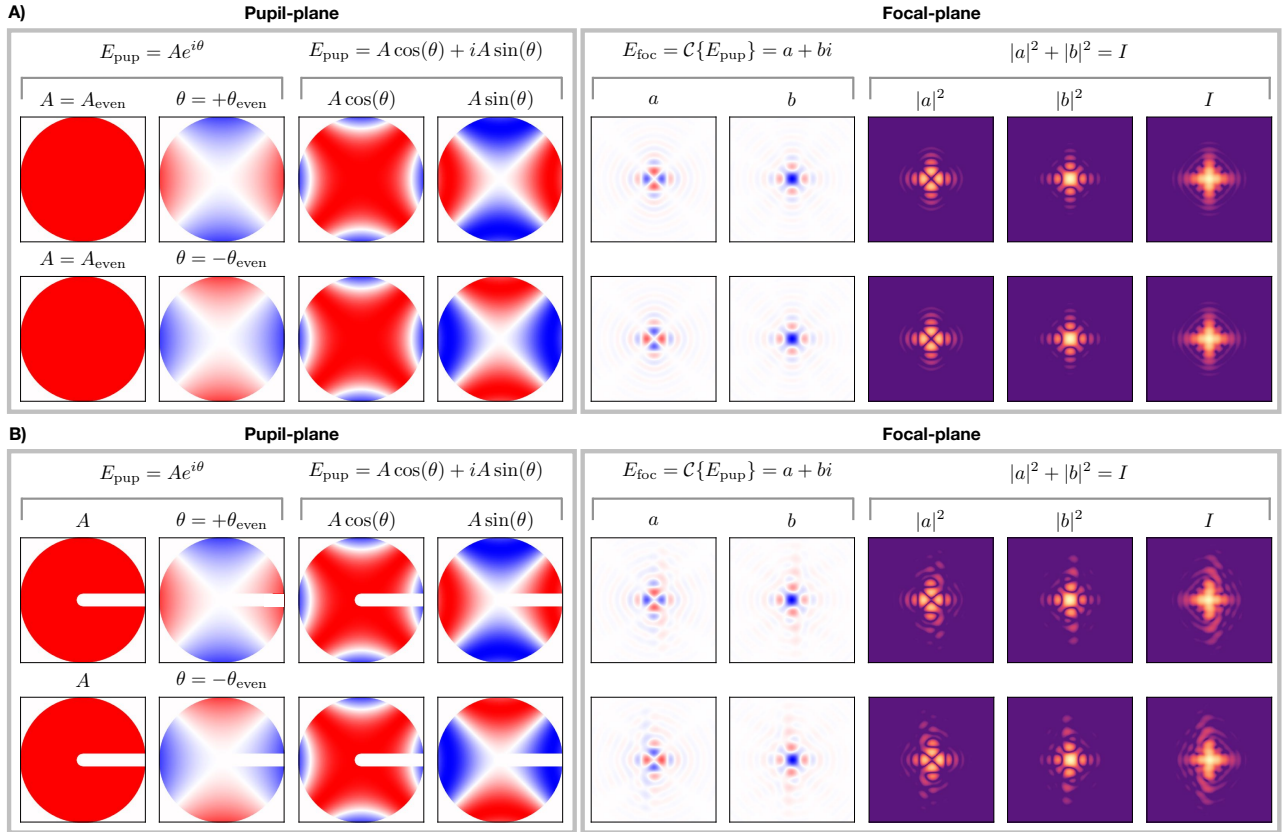


Fig. 2. Focal- and pupil-plane quantities of an even phase aberration (astigmatism) through a (a) symmetric and (b) asymmetric pupil. The two rows show opposite signs for the phase aberration. The columns in the pupil-plane box show (from left to right) the amplitude, phase, real and imaginary electric field. In the focal-plane box, the columns show the real and imaginary electric fields, the power in the real and imaginary electric fields, and the total power.

only the imaginary term contains sign information on the aberration. In this example, when they are propagated to the focal-plane, $A \cos(\theta)$ will go to ib and $iA \sin(\theta)$ to a (shown, respectively, in the first and third row of Table 1). These terms are still even and the sign information on the aberration is encoded in the real part of the focal-plane electric field. The resulting PSF recorded by the detector is even. These steps are shown in the top row of Figure 2 a. If the sign of the aberration flips (i.e a conjugation of the pupil-plane electric field), the PSF flips. But as the PSF is even, there is no morphology change recorded and thus the sign information cannot be retrieved. See the bottom row of Figure 2 a. One can only hope to determine the sign by measuring the real electric field, as that is where the sign information is encoded. For odd phase aberrations the sign flip will result in a morphology change and can therefore be measured.

If the same exercise is performed with a pupil amplitude

asymmetry, there will be a morphology change. This is shown in the top and bottom rows of Figure 2 b. The reason is that an even pupil amplitude, as shown in the top row of Figure 3 a, will only generate an imaginary electric field and therefore will not interfere with sign information containing real electric field caused by even aberrations. This is not the case for an asymmetric pupil amplitude, as it also generates a real electric field due to odd pupil amplitude (second row of Table 1). This real electric field will interfere with the aberration's real electric field and thus enable the sign determination. The electric field of an asymmetric pupil is shown in the bottom row of Figure 3 a. More examples on pupil-plane phase retrieval with focal-plane images can be found in Appendix A.

This is the working principle of the Asymmetric Pupil Fourier Wavefront Sensor (Martinache 2013). Interestingly, other focal-plane wavefront sensing methods such as the differ-

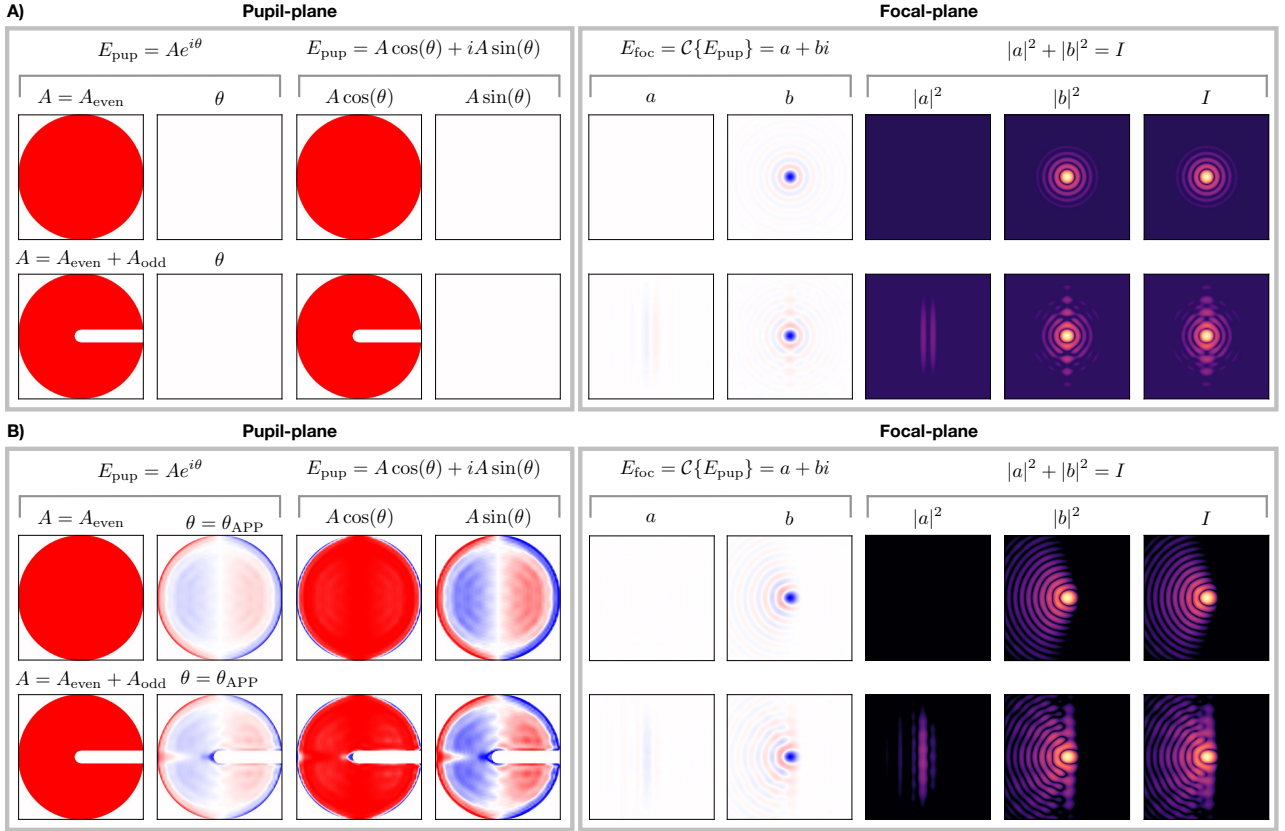


Fig. 3. Focal- and pupil-plane quantities for (a) a symmetric and asymmetric pupil, and (b) vAPPs designed for these pupils. The columns in the pupil-plane box show (from left to right) the amplitude, phase, real and imaginary electric field. In the focal-plane box, the columns show the real and imaginary electric fields, the power in the real and imaginary electric fields, and the total power.

ential Optical Transfer Function (Codona & Doble 2012) and Self-Coherent Camera (Baudoz et al. 2005) also rely on pupil asymmetries, albeit use other reconstruction algorithms. Classical phase diversity techniques (Gonsalves 1982; Paxman et al. 1992) come to a similar result by introducing a known, even phase aberration (e.g. defocus; odd phase modes can never be used, see the top row of Figure 3 b). This is the only other way of probing the real part of the focal-plane electric field (Table 1). Therefore, all these methods can now be understood as one family that probes the real focal-plane electric field by manipulating either pupil-plane phase or the pupil-plane amplitude. In the context of FPWFS with the vAPP, it is undesirable to use an even pupil-plane phase to break the sign ambiguity. This is because it fills up the dark-hole of the coronagraph and therefore prevents simultaneous science observations and wavefront measurements. On the other hand, vAPPs can be designed for pupils with an amplitude asymmetry, this will be shown in the next subsection, and therefore can combine science observations and wavefront sensing. This result also applies to other FPWFS techniques, for example see Appendix B how it affects spatial LDFC.

2.2. vAPP design for phase retrieval

Our framework of describing how pupil-plane phase and amplitude symmetries map to the focal-plane electric field also helps to understand some of the aspects of APP design and the requirements to turn an APP into a FPWFS. The optimization method that guarantees optimal APP designs, which turn out to be phase-only solutions, is detailed in Por (2017) and will not be discussed

here. The APPs considered in this section are all designed to have a one-sided D-shaped dark hole from $1.8 \lambda/D$ to $10 \lambda/D$ with a raw contrast of $< 10^{-5}$. The dark hole is defined as the area in the focal-plane where the raw contrast meets the requirement.

Suppose that we want to design an APP with a one-sided dark hole (= PSF with odd symmetry) for a symmetric aperture (A_{even}). Such an aperture will yield a completely imaginary focal-plane electric field with even symmetry as shown in the top row of Figure 3. As the APP manipulates phase, this imaginary electric field needs to be cancelled on one side using pupil-plane phase. The last row of Table 1 indicates that the only way of removing an odd part of the imaginary focal-plane electric field (ib) for an even aperture (A_{even}) is by introducing a purely odd pupil-plane phase. This indeed results from the optimization method, see the top row of Figure 3 b, and Figure 4 a.

An asymmetric aperture, e.g. as shown in the bottom row of Figure 3 a, yields a focal-plane electric field with real and imaginary components. The asymmetric aperture consists of a central obscuration that is 10% of the aperture diameter and a bar with the same width that connects the central obscuration with the edge of the aperture. Cancelling such a focal-plane electric field on one side to create the dark hole requires an odd pupil phase to cancel the imaginary focal-plane electric field (last row of Table 1), and an even pupil-plane phase for the real part (third row of Table 1). Again, this is indeed the solution the optimization method comes up with, as shown in the bottom rows of Figure 3 b, and Figure 4 b. Crucially for FPWFS with the APP, the real focal-plane electric field originating from the *odd* pupil amplitude component cannot be completely removed by the *even*

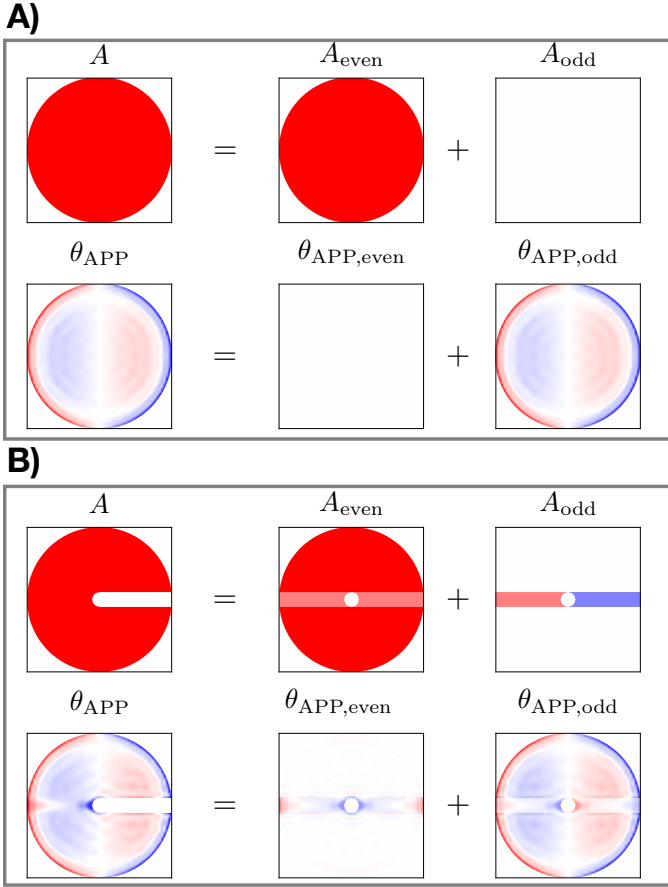


Fig. 4. Decomposition of the pupil-plane phase and amplitude into their even and odd constituents for the two APP designs in Figure 3 b. (a) An APP designed for a symmetric aperture. (b) An APP designed for an asymmetric aperture.

pupil-plane phase, but is enhanced on the bright side of the APP coronagraphic PSF, as shown in the bottom row of Figure 3 b.

To summarize, FPWFS capabilities of the vAPP are fundamentally enabled by the pupil amplitude asymmetry. This asymmetry introduces a real focal-plane electric field that interferes with the even pupil-plane phase aberrations and enables their sign determination. The vAPP allows for simultaneous science observations by removing the the real electric field from the dark hole and moving it to the bright field. See Appendix A for figures that demonstrate this.

3. Aberration estimation algorithm

3.1. Maximum a posteriori estimation

We have developed an algorithm that gives a maximum a posteriori (MAP) estimation of the phase aberrations by maximizing the posterior likelihood $p(\alpha, N_p, N_b, v, L|D)$. It takes into account a physical, non-linear model of the vAPP (shown in the next subsection), noise statistics and prior knowledge of the estimated parameters. Given an image D the algorithm estimates α , a vector containing the amplitudes of the phase aberration modes of interest. It has the option to additionally estimate the following parameters: the number of source photons in the image N_p , the background level N_b , the fractional degree of circular polarization v , and the fractional strength of the leakage PSF integrated over the spectral band L . These parameters are summarized in Table 2. The estimation of the phase aberrations is more conve-

niently performed by minimizing $\mathcal{L}(\alpha, N_p, N_b, v, L|D)$; the negative logarithm of the likelihood function ($\mathcal{L}(\alpha, N_p, N_b, v, L|D) = -\ln[p(\alpha, N_p, N_b, v, L|D)]$):

$$\hat{\alpha}, \hat{N}_p, \hat{N}_b, \hat{v}, \hat{L} = \arg \min_{\alpha, N_p, N_b, v, L} \mathcal{L}(\alpha, N_p, N_b, v, L|D). \quad (19)$$

Parameters of \mathcal{L} that have been estimated are denoted with a hat. The objective function \mathcal{L} is given by:

$$\mathcal{L}(\alpha, N_p, N_b, v, L|D) = \sum_x \frac{1}{2\sigma_n^2} (D - M(\alpha, N_p, N_b, v, L))^2 + \mathcal{R}(\alpha) \quad (20)$$

with $D(x)$ the 2D image containing the data and $M(\alpha, N_p, N_b, v, L)$ the 2D model of the vAPP PSF which will be detailed in the next subsection. The algorithm fits the model to the entire image and thus does not exclude any regions in $D(x)$. The sum over x is over all spatial pixels. The algorithm will be applied to long-exposure images that have high photon numbers, which can be approximated to contribute a spatially changing Gaussian noise (σ_p^2), and detector noise, assumed to be white Gaussian noise (σ_d^2). Therefore, the total noise is Gaussian and the variance is the sum of the two independent processes variances ($\sigma_n^2 = \sigma_p^2 + \sigma_d^2$). Prior information on the phase aberrations is taken into account explicitly by adding a term $\mathcal{R}(\alpha)$ to the objective function, which is given by:

$$\mathcal{R}(\alpha) = \frac{1}{2} \sum_{k=1}^N \frac{\alpha_k^2}{k^\gamma}, \quad (21)$$

with α_k the modal coefficients and γ the assumed power spectrum. This term penalizes high spatial frequency modes according to assumptions on or measurements of the spatial power spectrum. Implicitly, we also regularize by using a truncated mode basis.

We minimize \mathcal{L} using the bounded limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (L-BFGS-B; Byrd et al. 1995), which is a quasi-Newton optimization algorithm that assumes a differentiable scalar objective function. The algorithm accepts analytically calculated gradients to increase the speed and accuracy. Such analytical expressions of Equation 20's gradients are given in Appendix C. The algorithm also requires a start position for the estimated parameters. Generally, N_p, N_b, v and L can be easily estimated from the data, as will be shown in section 5. From experience, setting $\alpha = 0$ (i.e. no aberration) works best when there is no initial guess available for the aberrated wavefront.

3.2. Coronagraph model

The vAPP is a half-wave retarder with a spatially varying fast-axis; these spatial variations induce geometric phase on the circular polarization states. Opposite circular polarizations receive the opposite phase, creating two similar, but mirror-imaged coronagraphic PSFs, see Figure 1. Additionally, as the optic will not be perfectly half-wave, part of the light will not receive the desired phase and will create an extra non-coronagraphic PSF; we will refer to this PSF as the leakage. The relative intensity of the coronagraphic PSFs depends on the fractional degree of circular polarization v (normalized to the total intensity) and the amount of leakage. In the model we assume a single point source and therefore the image plane is given by:

$$M(\alpha, N_p, N_b, v, L) = N_p \left(\sum_{j=1}^3 a_j(v, L) I_{\text{foc},j}(\alpha) \right) + N_b, \quad (22)$$

Table 2. Parameters presented in section 3.

| Variable | Description |
|-----------------------|--|
| α | Vector containing the modal coefficients α_i . |
| $\hat{\alpha}$ | Estimation of α by algorithm. |
| γ | Assumed power spectrum in the regularization term. |
| θ_j | Pupil-plane phase of PSF j (aberration + vAPP). |
| θ_{APP} | Pupil-plane phase of the vAPP. |
| σ_d^2 | Variance of detector noise. |
| σ_n^2 | Total noise variance ($\sigma_n^2 = \sigma_p^2 + \sigma_d^2$). |
| σ_p^2 | Variance of photon noise. |
| ϕ_i | Pupil-plane phase of aberration mode i . |
| a_j | PSF j 's relative intensity, see Equations (23) to (25). |
| A | Pupil-plane amplitude. |
| $C\{\cdot\}$ | Fraunhofer propagator (Goodman 2005). |
| E_{pup} | Pupil-plane electric field. |
| D | Image used for wavefront sensing. |
| $I_{\text{foc},j}$ | PSF j , see Equation 26. |
| L | Fractional strength of leakage PSF. |
| \hat{L} | Estimation of L by algorithm. |
| \mathcal{L} | Objective function, see Equation 20. |
| M | PSF model of vAPP, see Equation 22. |
| N_p | Total number of photons in image. |
| \hat{N}_p | Estimation of N_p by algorithm. |
| N_b | Background level in image. |
| \hat{N}_b | Estimation of N_b by algorithm. |
| \mathcal{R} | Regularization term, see Equation 21. |
| ν | The fractional degree of circular polarization. |
| $\hat{\nu}$ | Estimation of ν by algorithm. |

with N_p the total number of photons in the image, N_b the background level (e.g. sky background, dark current or bias), $I_{\text{foc},j}$ the image from by either left- ($j=1$) or right-handed ($j=2$) circular polarization or leakage ($j=3$). The $a_j(\nu, L)$ terms describe the relative intensities of these PSFs and are given by:

$$a_1(\nu, L) = \frac{1 + \nu}{2}(1 - L), \quad (23)$$

$$a_2(\nu, L) = \frac{1 - \nu}{2}(1 - L), \quad (24)$$

$$a_3(\nu, L) = L. \quad (25)$$

PSFs are normalized such that the sum over all pixels in the image is one, i.e. $\sum_r I_{\text{foc},j} = 1$. Note that we can incoherently add the leakage term and the coronagraphic PSFs, because orthogonal polarization states emitted by blackbody sources are incoherent, even when they are transformed to the same polarization state (Fresnel-Arago laws; Kanseri et al. 2008; Mujat et al. 2004). The $I_{\text{foc},j}$ is given by:

$$I_{\text{foc},j}(\alpha) = |C\{E_{\text{pup}}(\alpha)\}|^2 \quad (26)$$

$$= |C\{Ae^{i(\theta_j + \sum_i \alpha_i \phi_i)}\}|^2 \quad (27)$$

with A the amplitude of the aperture, $\theta_j = \pm\theta_{\text{app}}$ for $j = 1, 2$ and $\theta_j = 0$ for $j = 3$. The phase aberration is expanded in the mode basis $\{\phi_i\}$, with α_i being the estimated coefficients. This last equation shows the non-linearity of the model: 1) we estimate phase, which is contained in the complex exponent, and 2) the estimation is performed on intensity images, which are the square of the electric field.

4. Simulations

In this section we explore the performance of the algorithm in idealised circumstances by numerical simulations. Performance will be reported in absolute units (nm) in the context of the SCEXAO vAPP and in relative units (fractional λ) to allow for comparison with other methods.

The design of the SCEXAO vAPP is detailed in Doelman et al. (2017); here we only discuss the most relevant details. The D-shaped dark holes have a raw contrast of 10^{-5} w.r.t. the peak flux of the coronagraphic PSF and extend from 2 to 11 λ/D . Besides the two coronagraphic PSFs, and the central leakage PSF, two additional phase diversity PSFs were added. These PSFs have opposite defocus aberration and can be used for classical phase diversity, but are not considered in this work. The top row of Figure 5 shows the pupil-plane amplitude for which the vAPP was designed, the subsequent phase design and the focal-plane intensity. As detailed in the bottom row of the figure, there is a clear amplitude asymmetry due to the blocking of dead DM actuators. This odd amplitude component A_{odd} , combined with the even phase θ_{even} that cancels the effect of the asymmetry in the dark holes, results in the real focal-plane electric field that will be used to probe even pupil-plane phase aberrations. Approximately 3.5 % of the focal plane power is in the real electric field and therefore available to sense even phase aberrations.

Simulations are performed with the HCIPy package (Por et al. 2018), an open-source software for high-contrast imaging simulations that is available on GitHub¹. The simulations are monochromatic ($\lambda=1550$ nm). They sample the pupil-plane with 256×256 pixels and use the Fraunhofer approximation to propagate the electric field to the focal-plane, which is sampled with 150×150 pixels for 3 pixels per λ/D . The aberration estimation algorithm is Python-based, uses the HCIPy package for the coronagraph model and the Scipy library (Jones et al. 2014) to minimize the objective function.

4.1. Photon and read noise sensitivity

Figure 6 presents the sensitivity of the algorithm and the SCEXAO vAPP when considering varying levels of photon and read noise. The modal basis set used for these tests consisted of the first 30 Zernike modes, starting at defocus. All induced aberrations were a linear sum of these modes. The aberrated phase screen is created by first calculating a phase screen following a spatial power spectrum with a slope of -2.5. This slope is steeper than a slope of -2, usually considered for the power spectrum of NCPA, but was chosen to put more power in the low-order modes as we focus on measuring those. Subsequently, this phase screen is projected onto the 30 Zernike modes, which were added together and scaled to a $\sim \lambda/16$ (100 nm) RMS WFE. The algorithm estimated the same Zernike modes as were used to generate the phase screen. For every photon number N_p ($10^3 - 10^9$ photons) we generated 10 different phase screens and, for each phase screen, 10 photon noise realizations. Furthermore, four levels of read noise were tested: 0 e⁻, 0.1 e⁻, 1 e⁻ and 10 e⁻. The results of the simulations are plotted in Figure 6, where the dots denote the mean RMS WFE per photon number and read noise level, and the error bars denote the 1σ deviation. The dotted line is a fit of a $\propto 1/\sqrt{N_p}$ function to the pure photon noise simulations, confirming that the performance scales with the photon noise. The algorithm therefore is photon noise limited, i.e. no algorithmic effects limit the performance. For the photon-limited case,

¹ <https://github.com/ehpor/hcipy>

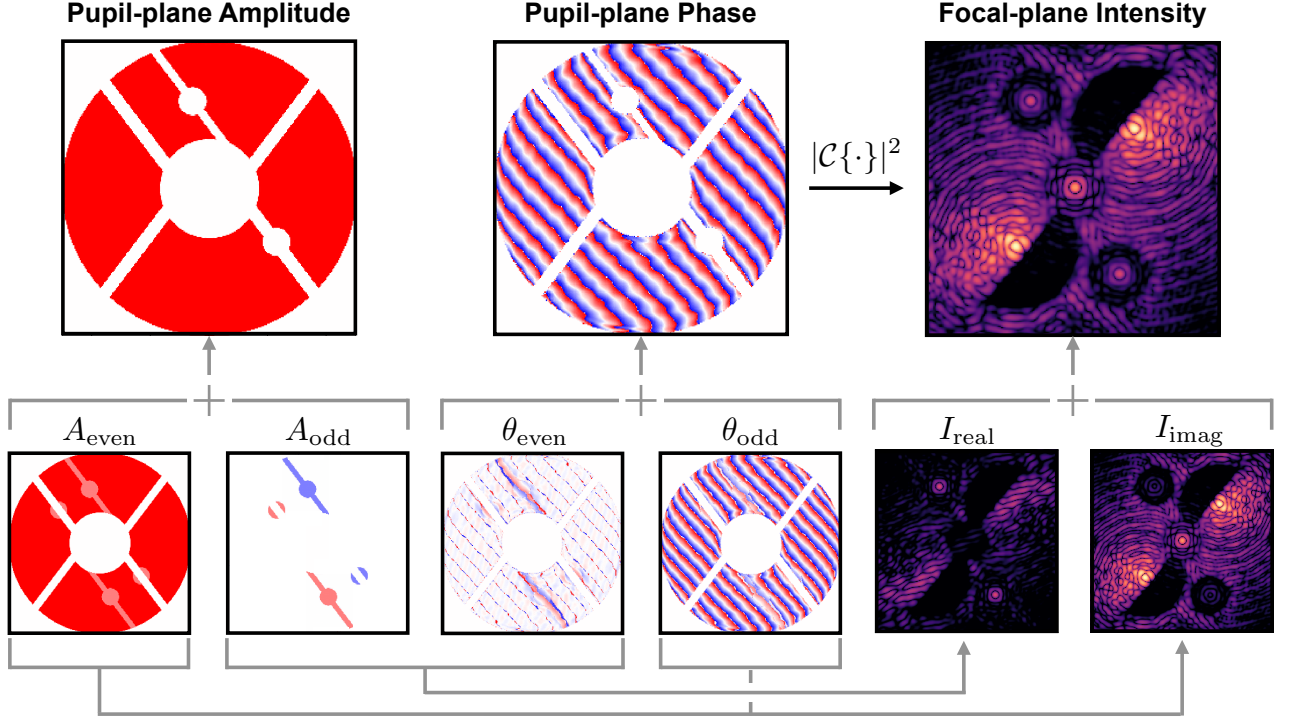


Fig. 5. Amplitude and phase design of the SCExAO vAPP (Doelman et al. 2017) and its resulting PSF. The pupil-plane quantities are decomposed into their even and odd constituents, while the focal-plane PSF is decomposed into the power in the real and imaginary part of the electric field.

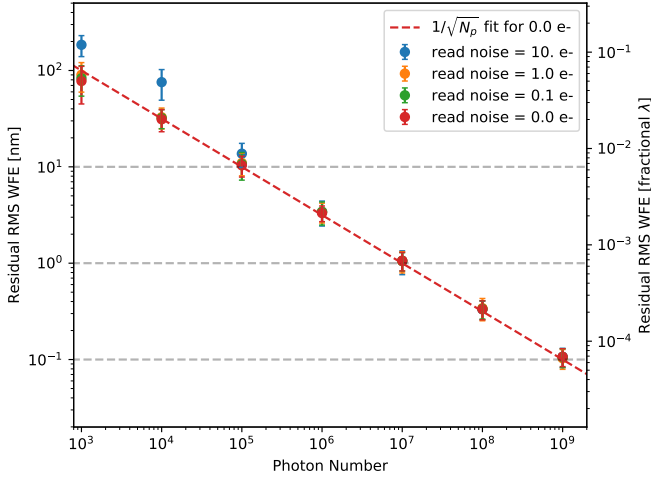


Fig. 6. Residual RMS WFE as a function of photon number N_p for the SCExAO vAPP. Different colours denote various read noise levels. The dotted line shows a $1/\sqrt{N_p}$ fit to the pure photon noise case. The RMS WFE in nm assumes $\lambda = 1550$ nm.

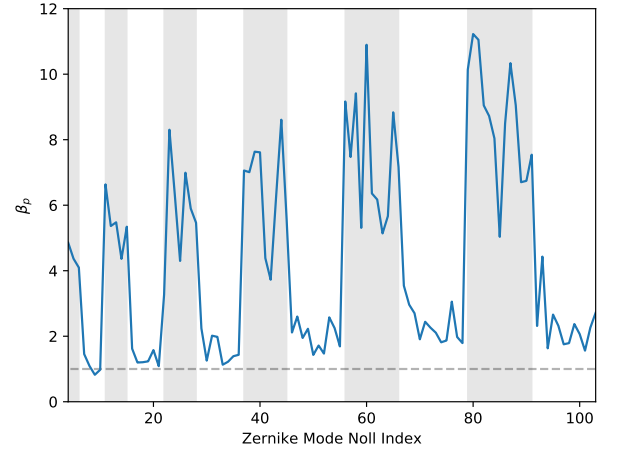


Fig. 7. Photon noise sensitivity of the SCExAO vAPP to the first 100 Zernike modes, starting at defocus. The grey boxes denote the even Zernike modes, and the white boxes the odd Zernike modes.

a $< \lambda/1000$ residual RMS WFE is reached with 10^7 photons. In the context of the SCExAO system (section 5) this would be a 1-nm residual WFE for a ~ 5 second integration on a $m_H = 8$ star. The simulations also show that increasing the read noise to 0.1 e^- or 1 e^- does not significantly impact the performance of the algorithm, i.e. the error is still dominated by photon noise. For a read noise level of 10 e^- and photon numbers $< 10^6$, the read noise starts to dominate the error and decreases the performance by a factor $\sim 2 - 3$.

4.2. Mode photon noise sensitivity

In Guyon (2005) the photon noise sensitivity β_p of a WFS to a single mode m is defined as:

$$\beta_p = \sigma_m \sqrt{N_p}, \quad (28)$$

with σ_m the reconstruction error for mode m and N_p the number of photons in the image. According to Guyon (2005), the ideal WFS will have $\beta_p = 1$ for all reconstructed modes. Here we explore the sensitivity of the SCExAO vAPP to the first 100 Zernike modes (starting with defocus as the vAPP is tip/tilt insensitive), expecting, as discussed in section 2 that the vAPP has

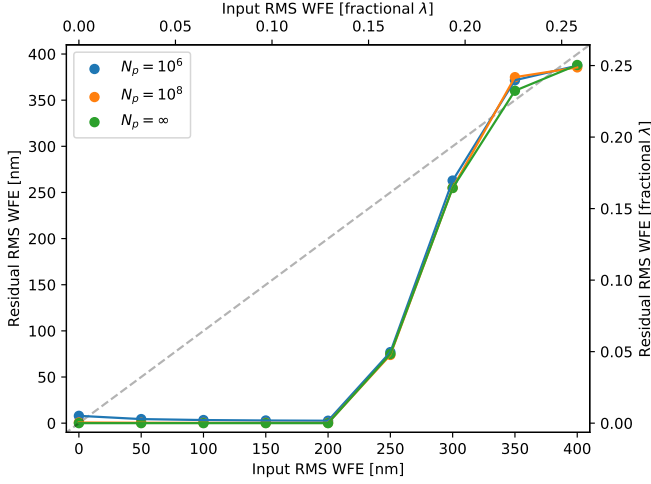


Fig. 8. Residual RMS WFE as function of introduced RMS WFE for the SCEExAO vAPP. The colours denote various photon numbers. The dashed, grey line separates the regions where the wavefront improves or deteriorates. The wavelength is 1550 nm.

a lower sensitivity (resulting in a higher β_p) for the even modes than for the odd modes. The reconstruction error σ_m was determined by calculating the standard deviation of the $\hat{\alpha}_m$ distribution, obtained after simulating 100 different photon noise realisations per photon number (10^7 , 10^8 and 10^9) and estimating only $\hat{\alpha}_m$ (no other modes were estimated simultaneously). This resulted in three β_p 's (for every tested photon number), which were subsequently averaged to obtain the final result. The results are shown in Figure 7, with the ideal WFS ($\beta_p = 1$) shown as the horizontal grey dotted line. The grey boxes denote the even modes, and the white boxes the odd modes. The first thing to note is that the even modes perform a factor $\sim 4 - 10$ worse than the odd modes. This is because, as discussed in section 2, the real part of the focal-plane electric field probes the even modes, and is in the case of the SCEExAO vAPP is relatively weak: $\sim 3.5\%$ of the total power is in the real electric field. Therefore, it is expected that β_p is $\sqrt{1/0.035} = 5.3$ times higher for even modes compared to the odd modes, which is approximately observed. Furthermore, it should be noted that the spatial frequency content of the Zernike modes increases with Noll index. Therefore, due to lower SNR at higher frequencies, sensitivity to both even and odd modes decreased with increasing mode number. The sensitivity to even modes decreases faster than the odd mode sensitivity, this is because the real part of the focal-plane electric field is limited in extent as compared to the imaginary part (shown in Figure 5). Note that this analysis does not include any cross-talk effects, which will be taken into account in the next section.

4.3. Dynamic range algorithm

Here we explore the maximum RMS WFE for which the algorithm can accurately recover estimates for each modal coefficient in a single step. This limitation is mainly driven by modal crosstalk and will determine the WFE for which the algorithm can converge to a low ($\sim \lambda/1000 - \lambda/100$ or 1-10 nm) residual RMS WFE in one iteration, assuming that the estimation is perfectly applied by the DM. We simulated phase screens consisting of 30 Zernike modes (starting with defocus), following a spatial power spectrum with a -2.5 slope, with increasing WFE.

For RMS WFE values between 0 and $\sim \lambda/4$ (400 nm) we simulated 10 wavefronts, with either no photon noise ($N_p = \infty$) or high photon numbers ($N_p = 10^6, 10^8$). For the simulations with photon noise, we simulated 10 photon-noise realisations per phase screen. The algorithm estimated the same Zernike modes. The residual RMS WFE was determined by subtracting the estimation from the initial, aberrated wavefront. The mean residual RMS WFEs are shown in Figure 8. The performance of the algorithm does not significantly vary between the different photon number realisations. The algorithm converges in one iteration with an input aberration RMS of up to $\sim \lambda/8$ or 200 nm. For WFE between $\lambda/8 - \lambda/5$ (200-300 nm), the algorithm still improves the wavefront, but there will be significant residual WFE after one iteration. In this regime, the algorithm will converge to $< \lambda/1000$ or 1-nm RMS WFE within two or three iterations in closed-loop operation. A WFE larger than $\lambda/5$ (300 nm) RMS does not significantly improve after one iteration. Note that this performance is very similar to the non-coronagraphic results reported by Martinache et al. (2016) for the Asymmetric Pupil Fourier Wavefront Sensor.

5. Demonstration at SCEExAO

5.1. SCEExAO

The Subaru Coronagraphic Extreme Adaptive Optics (SCEExAO) instrument (Jovanovic et al. 2015) is a high-contrast imaging instrument at the Subaru Telescope. It operates downstream of the AO188 system (Minowa et al. 2010), which gives an initial, low-order correction. SCEExAO drives a 2000-actuator deformable mirror (DM) based on wavefront measurements from a pyramid wavefront sensor (PYWFS) in the optical - NIR (600-900 nm) that can run at 3.5 kHz, but most often runs at 2 kHz. There are 45 actuators across the pupil, giving the system the ability to correct out to $22.5\lambda/D$. The real-time control of SCEExAO is handled by the compute and control for adaptive optics (cacao) package (Guyon et al. 2018). Cacao allows for NCPA corrections to be handled by a separate DM channel, where the software automatically sends offsets to the PYWFS. These offsets ensure that the PYWFS does not sense and attempt to control the NCPA corrections. The instrument routinely achieves Strehl ratios $> 80\%$ in H-band on-sky. SCEExAO feeds the post-AO JHK-bands to the integral field spectrograph CHARIS (Peters-Limbach et al. 2013; Groff et al. 2014). The ultimate goal is to operate the vAPP FPWFS with the CHARIS data; however, the tests were performed with the internal NIR camera for ease of operation. The NIR camera has recently (Lozi et al. 2018) been upgraded to a C-RED 2 detector (Feautrier et al. 2017), which has a 640×512 pixel InGaAs sensor that is cooled to -40°C .

5.2. Algorithm implementation in SCEExAO

The algorithm has to be adapted to the SCEExAO parameters and the vAPP design. To calibrate the model of the vAPP we had to take into account the following:

- **Pupil undersizing:** to account for pupil misalignments the design of the vAPP was undersized by 1.5% (Doelman et al. 2017) w.r.t. the nominal SCEExAO pupil.
- **Pupil rotation:** to optimally block dead actuators in the DM, the SCEExAO pupil was rotated by -6.25° , affecting the orientation of the PSF on the detector.
- **Detector pixel scale:** for accurate propagation to the focal-plane of the NIR camera the pixel scale was determined manually (which was deemed sufficient for low-order Zernike

mode measurement and correction) to be 15.6 mas per pixel at 1550 nm.

The coronagraph model described in section 3 assumes a monochromatic light source, and the closest available option, in terms of bandwidth, for the NIR camera in SCEXAO is the $\Delta\lambda = 25$ nm filter centered around $\lambda = 1550$ nm. For this filter, a 128×128 pixel subwindow in the NIR internal camera is sufficient to contain the vAPP PSFs. The image fed to the algorithm consists of the sum of a 1000 unsaturated images (either of the internal source or the star; individual exposure times are on the order of 1-10 milliseconds) that are dark-subtracted, aligned with a reference PSF generated with the coronagraph model and subsequently stacked and calibrated for the camera system gain ($2.33 \text{ e}^-/\text{ADU}$; Feautrier et al. 2017). The large number of stacked images should generally bring the algorithm into the 1-nm residual RMS WFE regime (Figure 6) in ideal circumstances and is the best performance we can expect. After image acquisition and before running the algorithm, the parameters N_p , N_b , v and L have to be estimated:

- N_b is estimated first by selecting the corners in the image where there is no light from the vAPP.
- N_p is estimated by subtracting N_b from the image and summing the values in all pixels.
- The grating in the vAPP acts as a circularly polarized beam-splitter, thus v is determined by aperture photometry on the two coronagraphic PSFs. From these intensity measurements, I_1 and I_2 , the fractional degree of circular polarization can be determined by:

$$v = \frac{I_1 - I_2}{I_1 + I_2} \quad (29)$$

- Similarly, the amount of leakage L is calculated by aperture photometry on the two coronagraphic (I_1 and I_2) and the leakage PSFs (I_3):

$$L = \frac{I_3}{I_1 + I_2 + I_3} \quad (30)$$

After this initial estimation, the estimation of these parameters can also be improved by the algorithm. This is generally not done for N_b , v and L as it has not been found to improve the wavefront estimation. On the other hand, we have found that estimating N_p using the algorithm does improve the estimation. Another concern is that estimating some of these parameters could induce cross-talk with the aberration coefficients. An example of such cross-talk would be between the wavefront aberration coma and v , as both parameters result in relative brightness changes between the two coronagraphic PSFs. Therefore, the algorithm could tune either parameter to fit a relative brightness change and arrive at the wrong result. We did not extensively study the effects of this cross-talk.

The closed-loop correction is performed by phase conjugation, where the DM command $\theta_{\text{DM},i}$ at iteration i is calculated as:

$$\theta_{\text{DM},i} = \theta_{\text{DM},i-1} - \frac{g}{2} \sum_j \alpha_j \phi_j, \quad (31)$$

with g the closed-loop gain that can be freely chosen, the factor $\frac{1}{2}$ to account for the reflective nature of the DM, α_j the estimated modal coefficients and $\{\phi_j\}$ the mode basis.

When estimating 30 Zernike modes with the SCEXAO implementation of the algorithm, the computation time per iteration

is $\sim 30 - 50$ seconds. Generally, the convergence time increased when sensing more modes and larger wavefront errors. The convergence time was mainly limited by the Python library versions installed at SCEXAO, and more fundamentally by the implementation in Python.

5.3. Internal source demonstration

We performed FPWFS tests with the SCEXAO vAPP using the internal source in November and December 2018. The goal was to demonstrate the principle of FPWFS with the vAPP coronagraphic PSFs, and we therefore focussed on measuring and correcting low-order modes (30-50 Zernike modes, starting with defocus). Figure 9 shows these results, where the 1σ raw contrast was determined by calculating the standard deviation of $1 \lambda/D$ wide annuli covering both PSFs that were normalized by the maximum number of counts in the image. The top row in the figure presents the pre- and post-correction PSFs and the radial averaged contrast improvements (the subfigures a, b and c, respectively) for the nominal SCEXAO system using 30 Zernike modes, after one iteration with the algorithm ($g = 1$). The estimated wavefront is shown in Figure 10 a, and the measured WFE in the 30 Zernike modes is ~ 100 nm RMS. The dominant mode is astigmatism at ~ 86 nm RMS. The next three largest aberrations were coma, quadrafoil and secondary astigmatism at respectively ~ 28 nm, ~ 24 nm and ~ 23 nm RMS. Qualitatively, the wavefront improvement can be observed in the more symmetric leakage and phase diversity PSFs, and the dark hole is better defined as aberrated structure at a few λ/D has been removed. Quantitatively, the peak flux of the leakage increased by $\sim 10\%$, and the raw contrast at $2.5 \lambda/D$ improved from $\sim 2 \cdot 10^{-3}$ to $\sim 6 \cdot 10^{-4}$ (solid lines in Figure 9 c). Only the lowest Zernike modes (both even and odd) were corrected, which is clearly visible as the contrast improvement decreases with λ/D . The dashed-dotted lines in Figure 9 c show the results of a similar correction performed at another moment in time, but here the images used for the raw contrast determination were saturated such that the raw contrast was speckle-noise limited, instead of read noise limited as with the solid lines. The bottom row shows the result of a closed-loop correction where we introduced and estimated a 300 nm RMS WFE with 50 Zernike modes, which should be correctable according to Figure 8. Images b and c show the pre- and post-correction after 5 iterations of closed loop correction. The total duration of these corrections was around 4 minutes. These figures show that the algorithm can correct large WFE, recovering a PSF that is qualitatively very similar to what is shown in Figure 9 b. As shown in Figure 9 f, we recover similar or deeper raw contrasts compared to the nominal system after correction, Figure 9 c. In Figure 10 b and c, the introduced and estimated wavefronts are shown. The WFE for these figures are respectively ~ 293 nm and ~ 311 nm. Note that the estimated wavefront also includes WFE already present in the system before adding the known aberration, as this was not corrected before the test. Tests conducted on another day suggest that the remaining WFE after correction is on the order of ~ 59 nm RMS. These tests consist of sequential measurements of the nominal wavefront and a 150 nm RMS introduced wavefront error. These measurements were then compared to find a residual WFE. This error consists of evolved NCPA between measurements, DM calibration errors and algorithm errors. We did not investigate the amplitude of the individual error terms.

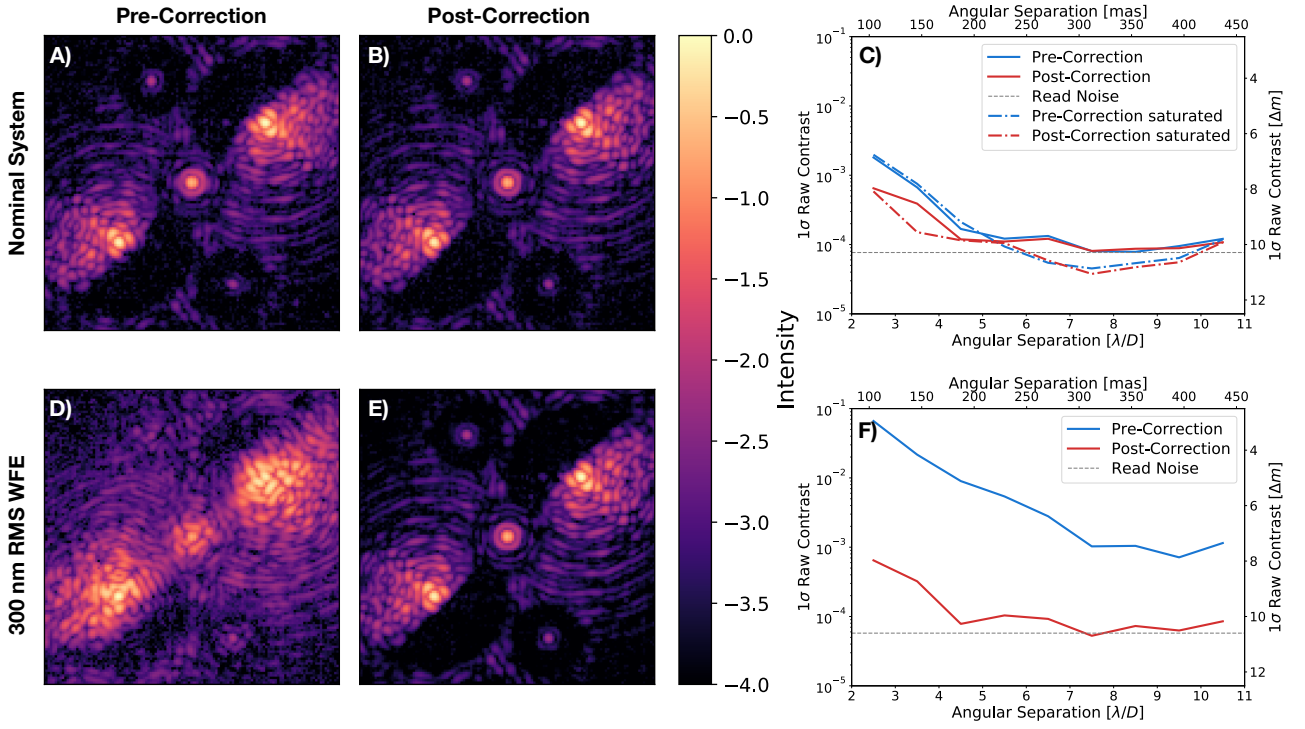


Fig. 9. Focal-plane images and 1σ raw contrast curves pre- and post-correction with the internal source within SCEXAO. (a) Pre- and (b) post-correction images for the nominal SCEXAO system. (c) The 1σ raw contrast curves for the correction of the nominal system. The solid lines relate to subfigures (a) and (b) and are read noise limited beyond $7.5 \lambda/D$, which is shown by the horizontal dashed line. The dot-dashed lines show a similar correction performed at a different time. For these lines the raw contrast determination were saturated such that they were speckle noise limited (instead of read noise limited), reaching deeper contrasts in the dark holes. (d) Pre- and (e) post-correction images for an introduced wavefront of 300 nm RMS. (f) The 1σ raw contrast curves for the correction of the system with the introduced 300 nm RMS wavefront, with again the horizontal dashed line denoting the read noise.

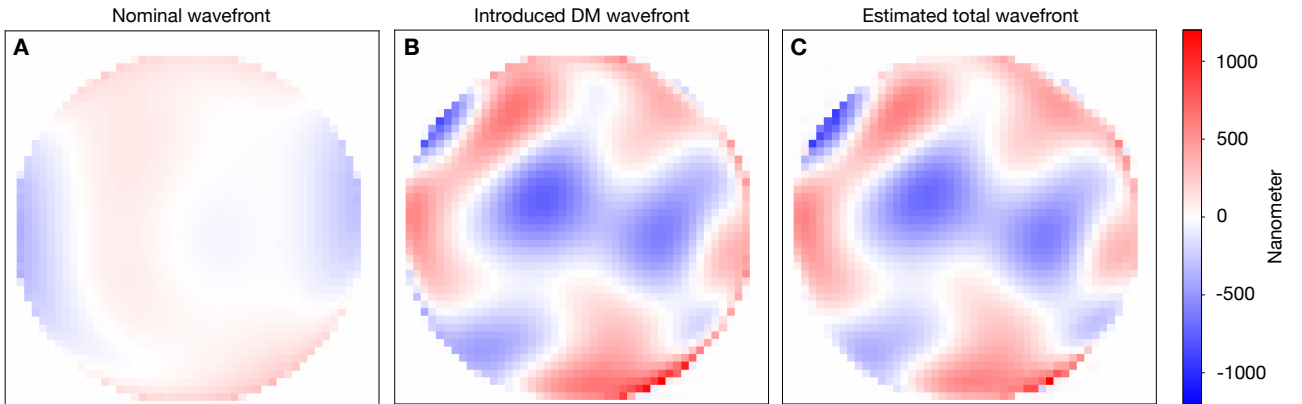


Fig. 10. (a) Estimated wavefront of the nominal SCEXAO system using the internal source and reconstructed by combining the estimated modal coefficients of the 30 Zernike modes. The estimated WFE is ~ 100 nm RMS. (b) The introduced and (c) estimated wavefront from closed-loop tests with the internal source. The introduced wavefront (b), consisting of 50 Zernike modes, has a RMS WFE of ~ 293 nm. The estimated wavefront (c), using the same 50 Zernike modes, with an RMS WFE of ~ 311 nm. Note that the estimated wavefront also includes other WFE already present in the system.

5.4. On-sky demonstration

On the 11th of January 2019 we observed Regulus ($m_H = 1.57$; Ducati 2002) during a SCEXAO engineering night to test the vAPP FPWFS on-sky, to evaluate the performance with an incoherent background due to residual turbulence and long exposure images, and to demonstrate with the true end-to-end optical system. Observing conditions were good - the seeing was $\sim 0.2'' - 0.4''$ - and predictive control was running (Guyon & Males 2017) during the tests. We conducted similar tests as with

the internal source of which the results are shown in Figure 11. The top row shows the results after 5 iterations of closed-loop with a gain=0.5 with the nominal SCEXAO system when estimating 30 Zernike modes. The 5 iterations took approximately 4 minutes. Qualitatively, the first Airy ring of the leakage PSF becomes more rounded and symmetric, the coronagraphic PSFs are more similar, the edge of the dark hole becomes better defined and some speckles in the dark hole are removed. The gain in peak flux of the leakage PSF is 6%. Shown in Figure 11 c, the contrast gain is moderate, improving by a factor ~ 2 between

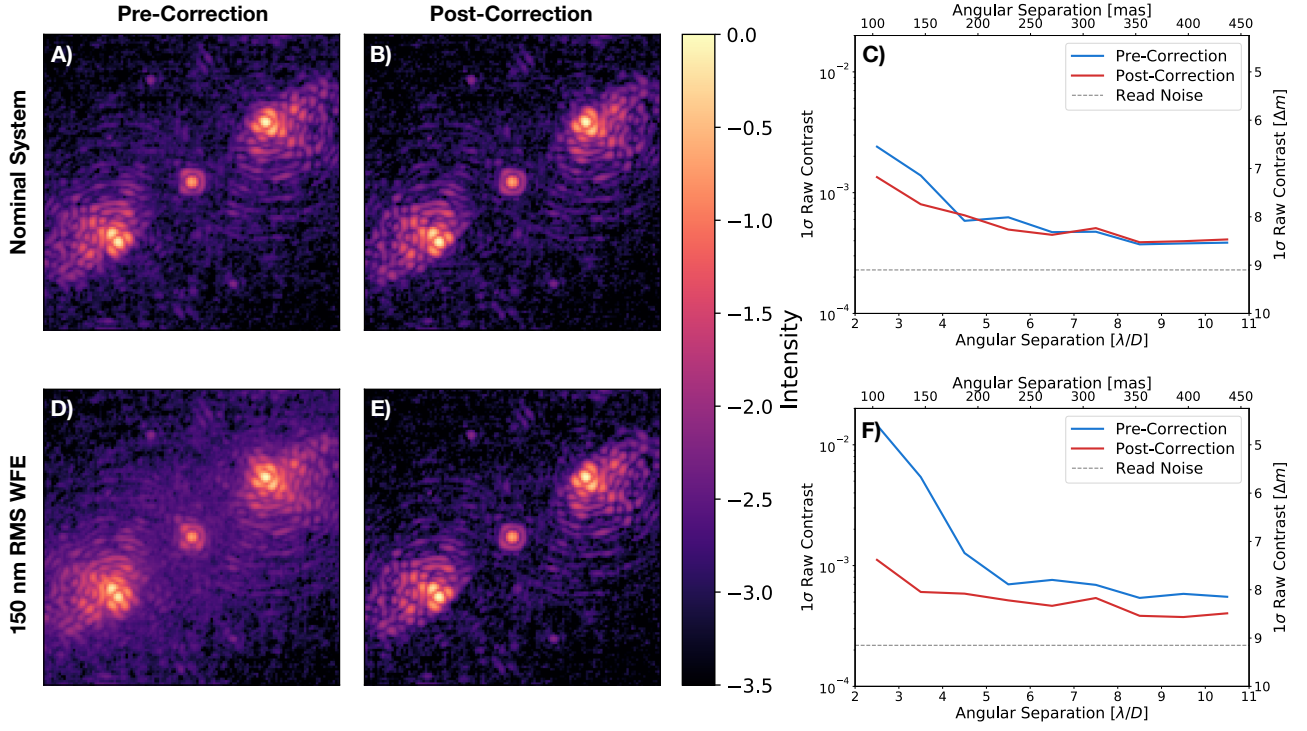


Fig. 11. Focal-plane images and 1σ raw contrast curves pre- and post-correction during on-sky observations. (a) Pre- and (b) post-correction images for the nominal SCExAO system. (c) The 1σ raw contrast curves for the correction of the nominal system. The horizontal dashed line denotes the read noise level. (d) Pre- and (e) post-correction images for an additional wavefront error of 150 nm RMS. (f) The 1σ raw contrast curves for the correction of the system with the additional 150 nm RMS wavefront error, with again the horizontal dashed line denoting the read noise.

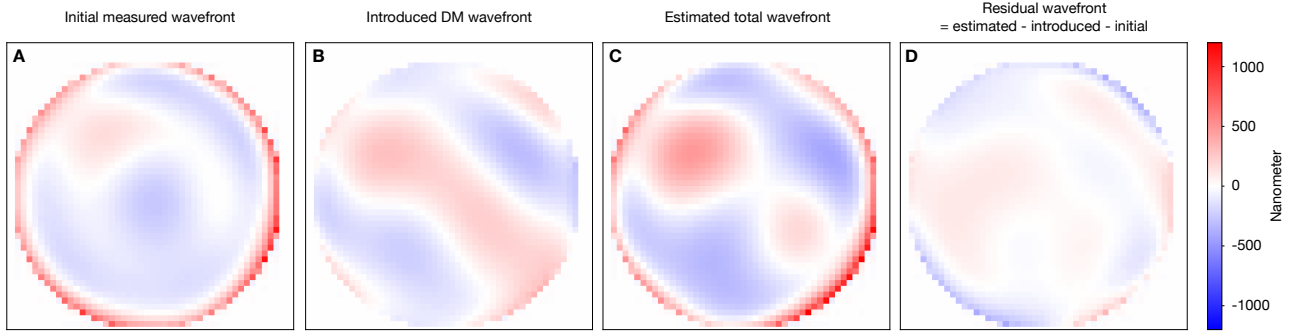


Fig. 12. (a) Estimated wavefront of the nominal SCExAO system during on-sky observations and reconstructed by combining the estimated modal coefficients of the 30 Zernike modes. The estimated WFE is ~ 187 nm RMS. The (b) introduced, (c) estimated and (d) residual wavefront from closed-loop tests during on-sky observations. (b) The introduced wavefront, consisting of 30 Zernike modes, with an RMS WFE of ~ 146 nm. (c) The estimated wavefront, using the same 30 Zernike modes, with an RMS WFE of ~ 244 nm. (d) The residual wavefront (introduced minus estimated and nominal wavefronts), the RMS WFE is ~ 78 nm.

2.5 and $4 \lambda/D$. The estimated wavefront is shown in Figure 12 a and the WFE is ~ 187 nm RMS. A notable structure in this estimated wavefront is the sharp increase in phase at the edges of the pupil. It is thought that this phase ring is a relatively static and real structure and originates from the upstream AO188 system. This is because the PYWFS reference offset contains a similar structure as well. The bottom row shows the correction of low order quasi-static errors from the instrument on-sky plus and additional 150 nm RMS WFE that was added to the DM consisting of 30 Zernike modes. After 5 iterations (gain = 0.5), shown in Figure 11 e, the algorithm recovers a similar PSF in a similar time as for the correction of the nominal system (Figure 11 c), both in morphology and in achieved raw contrast, which is shown in Figure 11 f. In Figure 12 b-d the introduced, estimated

and residual wavefronts are shown, with the WFE being respectively ~ 146 nm, ~ 244 nm and ~ 78 nm RMS. The residual wavefront was determined by subtracting the introduced and initial wavefronts from the estimated wavefront. This gives an estimate of the remaining wavefront error in the system. As discussed with the internal source results, the residual wavefront consists of evolved NCPA between measurements, DM calibration errors and algorithm errors. Note that the reported raw contrasts in Figure 11 c and f are worse than in Figure 9 c and f due to the incoherent speckle background created by the uncorrected, residual atmospheric wavefront errors.

6. Discussion and Conclusion

We have shown in section 2 that an asymmetry in the pupil amplitude enables the vAPP coronagraphic PSFs to measure both even and odd pupil phase modes, generalizing the Asymmetric Pupil Fourier Wavefront Sensor (APF-WFS; Martinache 2013) with spatial Linear Dark Field Control (LDFC; Miller et al. 2017) in the vAPP. The physical model for non-linear wavefront estimation developed in section 3 was tested in idealised simulations (section 4), confirming that the vAPP currently installed at SCExAO can sense the even modes, but is more sensitive to odd modes. Simulations suggest that the algorithm should be able to reach one-shot correction resulting in a $< \lambda/1000$ nm RMS WFE with approximately 10^7 photons, and that with one iteration it should be able to correct up to $\sim \lambda/8$ of RMS WFE. In section 5, we have demonstrated the principle with the vAPP in SCExAO, both with the internal source and on-sky, by measuring and controlling the 30 lowest Zernike modes and improving the raw contrast between 2.5 and $4 \lambda/D$ by a factor ~ 2 . Furthermore, we have shown that the algorithm can correct WFE's of 300 nm and 150 nm RMS with the internal source and on-sky, respectively. Although the contrast gains, both with the internal source and on-sky, are moderate, it does demonstrate that the coronagraphic PSFs of the vAPP can be used for wavefront sensing.

The FPWFS performance of the SCExAO vAPP can be improved in various ways: 1) better model calibration to increase the contrast gain, 2) measuring more and higher spatial frequency modes to increase the area of correction, and 3) improving the convergence speed of the algorithm. For this demonstration the model was calibrated ad-hoc by tuning the pupil undersizing, rotation and the detector pixel scale by hand. These parameters should be more accurately measured or be part of the fitted parameters. To measure more and higher spatial frequency modes we want to replace the current Zernike mode basis with a pixel-mode basis (Paul et al. 2013a). This should also reduce the effect of unsensed modes on the estimation, when estimating a truncated mode basis (Paul et al. 2013b). Improving the convergence speed of the algorithm would result in a higher correction cadence or the correction of a larger mode basis in a similar time. A relatively simple improvement would be using the multiprocessing capabilities in the latest python libraries not yet available on SCExAO, this would enable multi-core processing and give a maximum factor of 40 speed improvement. Theoretically, the best possible performance we can then expect is a wavefront estimation with sub 1 nm WFE every six seconds for targets down to a $m_H = 8$ (five seconds of integration and one second for estimation). Other coronagraph model improvements include the integral field spectrograph CHARIS, an accurate model of the focal-plane filter and the incoherent background due to residual turbulence effects. This is because science observations with the SCExAO vAPP are done with CHARIS. To prevent detector saturation due to bright fields of the vAPP, it requires a focal-plane neutral density filter that attenuates the bright fields. Therefore, exposures with the focal-plane filter in place will have a larger dynamic range. This increase in dynamic range will make the incoherent background, due to uncorrected atmospheric turbulence, a more prominent feature in the image compared to the current on-sky results. Thus, the incoherent background needs to be included within the coronagraph model as well (Herscovici-Schiller et al. 2017), also including any asymmetries in this background (Cantalloube et al. 2018).

In this article we have focused on a FPWFS based on a grating-vAPP. We operated it only in narrowband mode or in

an integral field spectrograph. Therefore, the FPWFS algorithm does not have to operate over broad wavelength ranges and the current coronagraph model would suffice. However, there are broadband imaging vAPPs implementations foreseen, as is detailed in Bos et al. (2018). These coronagraphs are more complicated in their optical design and therefore require a more advanced coronagraph model. For such broadband coronagraphs, the coronagraph model also needs to be extended to handle broadband FPWFS to take full advantage of the coronagraph. This can be done by evaluating the model at multiple wavelengths (Seldin et al. 2000) and would come at an increased computational cost. Operating in a broadband mode would increase the sensitivity as there is more light available, but speckles at larger λ/D will be washed out and therefore higher frequency aberrations will be harder to measure.

In this paper we discussed only FPWFS for vAPPs that generate asymmetric dark holes. But FPWFS with pupil-plane coronagraphs that create symmetric dark holes such as the 360° vAPP (Otten et al. 2014) and the shaped pupil (Kasdin et al. 2007) are desirable as well. Optimal designs of vAPPs with symmetric dark holes consist of only 0 and π phase structures (Por 2017). This results in a purely real pupil-plane electric field (Equation 2) and thus if the aperture is symmetric the focal-plane electric field of such a vAPP will be completely imaginary (Table 1) and does not support a FPWFS. Therefore, similar to one-sided dark holes, FPWFSing is enabled when an amplitude asymmetry is introduced. A similar argument can be given for shaped pupil coronagraphs, as these manipulate pupil-plane amplitude and thus only have a real pupil-plane electric field. A more important difference with one-sided dark hole PSFs is that such coronagraphs will not have a bright field in the other polarization that covers the dark hole of the coronagraph and thus does not directly probe the region of interest.

Future work will investigate optimization of the aperture asymmetry to enhance the wavefront sensing capabilities of the coronagraph. This also has implications for the APF-WFS and spatial LDFC controlling dark holes dug with other methods than the vAPP. With the current implementation of the vAPP and algorithm we can only sense phase aberrations. Uncorrected amplitude errors due to the atmosphere and instrumental errors will limit the raw contrast to $\sim 10^{-5}$ (Guyon 2018). This is at a level that is not yet reached by the vAPP and therefore, sensing only phase aberrations is sufficient for now. When this raw contrast is met it would be necessary to sense both pupil amplitude and phase aberrations. It is expected that the two coronagraphic PSFs alone do not contain sufficient diversity to estimate both in one image. For the current SCExAO vAPP design, an additional classical phase diversity image will likely provide the required diversity to enable amplitude estimation (Herscovici-Schiller et al. 2018), but will lower the science duty cycle. A better solution for future FPWFS vAPPs would be the addition of two strong phase diversity holograms, this would keep the science duty cycle at 100% at the cost of science throughput. Another use of FPWFS with the vAPP could be the fine co-phasing of multi-mirror telescopes from the image plane (Pope et al. 2014), such as for the upcoming Giant Magellan Telescope.

Integration of FPWFS with coronagraphy is a crucial step in the system wide integration of all optical modalities (e.g. spectroscopy and polarimetry) to get the best possible performance of high contrast imaging instruments. Major advantages are that NCPA can be measured up to the science focal-plane, and a 100% science duty cycle as science observations do not have to be interrupted to probe the dark hole. The vAPP has now shown

to be able to combine both FPWFS and coronagraphy in one optic.

Acknowledgements. The authors thank the referee for comments on the manuscript that significantly improved the presentation of the work. The authors warmly thank S.Y. Haffert and E.H. Por for useful discussion on focal-plane wavefront sensing. The research of Steven P. Bos, David S. Doelman, and Frans Snik leading to these results has received funding from the European Research Council under ERC Starting Grant agreement 678194 (FALCONER). The development of SCEXAO was supported by the JSPS (Grant-in-Aid for Research #23340051, #26220704, #23103002), the Astrobiology Center (ABC) of the National Institutes of Natural Sciences, Japan, the Mt Cuba Foundation and the directors contingency fund at Subaru Telescope. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. This research made use of HCIPy, an open-source object-oriented framework written in Python for performing end-to-end simulations of high-contrast imaging instruments (Por et al. 2018). This research used the following Python libraries: Scipy (Jones et al. 2014), Numpy (Walt et al. 2011) and Matplotlib (Hunter 2007).

References

- Baudoz, P., Boccaletti, A., Baudrand, J., & Rouan, D. 2005, *Proceedings of the International Astronomical Union*, 1, 553
- Berry, M. V. 1987, *Journal of Modern Optics*, 34, 1401
- Beuzit, J.-L., Vigan, A., Mouillet, D., et al. 2019, arXiv preprint arXiv:1902.04080
- Boehle, A., Glauser, A. M., Kenworthy, M. A., et al. 2018, in *Ground-based and Airborne Instrumentation for Astronomy VII*, Vol. 10702, International Society for Optics and Photonics, 107023Y
- Bordé, P. J. & Traub, W. A. 2006, *The Astrophysical Journal*, 638, 488
- Bos, S. P., Doelman, D. S., de Boer, J., et al. 2018, in *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation III*, Vol. 10706, International Society for Optics and Photonics, 107065M
- Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. 1995, *SIAM Journal on Scientific Computing*, 16, 1190
- Cantalloube, F., Por, E., Dohlen, K., et al. 2018, *Astronomy & Astrophysics*, 620, L10
- Close, L. M., Males, J. R., Durney, O., et al. 2018, arXiv preprint arXiv:1807.04311
- Codona, J. L. & Doble, N. 2012, in *Adaptive Optics Systems III*, Vol. 8447, International Society for Optics and Photonics, 84476R
- Côté, O., Allain, G., Brousseau, D., et al. 2018, in *Ground-based and Airborne Instrumentation for Astronomy VII*, Vol. 10702, International Society for Optics and Photonics, 1070248
- Davies, R., Alves, J., Clénet, Y., et al. 2018, in *Ground-based and Airborne Instrumentation for Astronomy VII*, Vol. 10702, International Society for Optics and Photonics, 107021S
- De Kok, R., Stam, D., & Karalidi, T. 2011, *The Astrophysical Journal*, 741, 59
- Doelman, D. S., Auer, F. F., Escuti, M. J., & Snik, F. 2019, *Optics letters*, 44, 17
- Doelman, D. S., Snik, F., Warriner, N. Z., & Escuti, M. J. 2017, in *Techniques and Instrumentation for Detection of Exoplanets VIII*, Vol. 10400, International Society for Optics and Photonics, 104000U
- Ducati, J. 2002, *VizieR Online Data Catalog*, 2237
- Feautrier, P., Gach, J.-L., Greffe, T., et al. 2017, in *Image Sensing Technologies: Materials, Devices, Systems, and Applications IV*, Vol. 10209, International Society for Optics and Photonics, 102090G
- Galicher, R., Baudoz, P., & Rousset, G. 2008, *Astronomy & Astrophysics*, 488, L9
- Gonsalves, R. A. 1982, *Optical Engineering*, 21, 215829
- Goodman, J. W. 2005, *Introduction to Fourier optics* (Roberts and Company Publishers)
- Groff, T. D., Kasdin, N. J., Limbach, M. A., et al. 2014, in *Ground-based and Airborne Instrumentation for Astronomy V*, Vol. 9147, International Society for Optics and Photonics, 91471W
- Groff, T. D., Riggs, A. E., Kern, B., & Kasdin, N. J. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 2, 011009
- Guyon, O. 2005, *The Astrophysical Journal*, 629, 592
- Guyon, O. 2018, *Annual Review of Astronomy and Astrophysics*, 56, 315
- Guyon, O. & Males, J. 2017, arXiv preprint arXiv:1707.00570
- Guyon, O., Miller, K., Males, J., Belikov, R., & Kern, B. 2017, arXiv preprint arXiv:1706.07377
- Guyon, O., Sevin, A., Ltaief, H., et al. 2018, in *Adaptive Optics Systems VI*, Vol. 10703, International Society for Optics and Photonics, 107031E
- Haffert, S., Bohn, A., de Boer, J., et al. 2019, *Nature Astronomy*, 1, 1
- Haffert, S., Wilby, M., Keller, C., et al. 2018, in *Adaptive Optics Systems VI*, Vol. 10703, International Society for Optics and Photonics, 1070323
- Herscovici-Schiller, O., Mugnier, L. M., Baudoz, P., et al. 2018, *Astronomy & Astrophysics*, 614, A142
- Herscovici-Schiller, O., Mugnier, L. M., & Sauvage, J.-F. 2017, *Monthly Notices of the Royal Astronomical Society: Letters*, 467, L105
- Hoijmakers, H., Schwarzs, H., Snellen, I., et al. 2018, *Astronomy & Astrophysics*, 617, A144
- Hunter, J. D. 2007, *Computing In Science & Engineering*, 9, 90
- Jones, E., Oliphant, T., & Peterson, P. 2014
- Jovanovic, N., Absil, O., Baudoz, P., et al. 2018, in *Proc. SPIE*, Vol. 10703, Adaptive Optics Systems VI
- Jovanovic, N., Martinache, F., Guyon, O., et al. 2015, *Publications of the Astronomical Society of the Pacific*, 127, 890
- Kanseri, B., Bisht, N. S., Kandpal, H., & Rath, S. 2008, *American Journal of Physics*, 76, 39
- Kasdin, N. J., Vanderbei, R. J., & Belikov, R. 2007, *Comptes Rendus Physique*, 8, 312
- Kenworthy, M. A., Absil, O., Carlomagno, B., et al. 2018, in *Ground-based and Airborne Instrumentation for Astronomy VII*, Vol. 10702, International Society for Optics and Photonics, 10702A3
- Lozi, J., Guyon, O., & et al. 2018, in *Adaptive Optics Systems VI*, Vol. 10703, International Society for Optics and Photonics
- Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, *Proceedings of the National Academy of Sciences*, 111, 12661
- Males, J. R., Close, L. M., Miller, K., et al. 2018, in *Adaptive Optics Systems VI*, Vol. 10703, International Society for Optics and Photonics, 1070309
- Marois, C., Lafreniere, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, *The Astrophysical Journal*, 641, 556
- Martinache, F. 2013, *Publications of the Astronomical Society of the Pacific*, 125, 422
- Martinache, F., Jovanovic, N., & Guyon, O. 2016, *Astronomy & Astrophysics*, 593, A33
- Martinez, P., Kasper, M., Costille, A., et al. 2013, *Astronomy & Astrophysics*, 554, A41
- Martinez, P., Loose, C., Carpentier, E. A., & Kasper, M. 2012, *Astronomy & Astrophysics*, 541, A136
- Mazoyer, J., Baudoz, P., Galicher, R., Mas, M., & Rousset, G. 2013, *Astronomy & Astrophysics*, 557, A9
- Miller, K., Guyon, O., & Males, J. 2017, *Journal of Astronomical Telescopes, Instruments, and Systems*, 3, 049002
- Miller, K., Males, J. R., Guyon, O., et al. 2018, in *Adaptive Optics Systems VI*, Vol. 10703, International Society for Optics and Photonics, 107031T
- Miller, K. L. 2018, PhD thesis, The University of Arizona
- Minowa, Y., Hayano, Y., Oya, S., et al. 2010, in *Adaptive Optics Systems II*, Vol. 7736, International Society for Optics and Photonics, 77363N
- Mujat, M., Dogariu, A., & Wolf, E. 2004, *JOSA A*, 21, 2414
- N'Diaye, M., Dohlen, K., Fusco, T., & Paul, B. 2013, *Astronomy & Astrophysics*, 555, A94
- N'Diaye, M., Martinache, F., Jovanovic, N., et al. 2018, *Astronomy & Astrophysics*, 610, A18
- N'Diaye, M., Vigan, A., Dohlen, K., et al. 2016, in *Adaptive Optics Systems V*, Vol. 9909, International Society for Optics and Photonics, 99096S
- Oh, C. & Escuti, M. J. 2008, *Optics letters*, 33, 2287
- Otten, G. P., Snik, F., Kenworthy, M. A., et al. 2017, *The Astrophysical Journal*, 834, 175
- Otten, G. P., Snik, F., Kenworthy, M. A., et al. 2014, in *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation*, Vol. 9151, International Society for Optics and Photonics, 91511R
- Pancharatnam, S. 1956, in *Proceedings of the Indian Academy of Sciences-Section A*, Vol. 44, Springer, 398–417
- Paul, B., Mugnier, L., Sauvage, J.-F., Dohlen, K., & Ferrari, M. 2013a, *Optics Express*, 21, 31751
- Paul, B., Sauvage, J.-F., & Mugnier, L. 2013b, *Astronomy & Astrophysics*, 552, A48
- Paul, B., Sauvage, J.-F., Mugnier, L., et al. 2014, *Astronomy & Astrophysics*, 572, A32
- Paxman, R. G., Schulz, T. J., & Fienup, J. R. 1992, *JOSA A*, 9, 1072
- Peters-Limbach, M. A., Groff, T. D., Kasdin, N. J., et al. 2013, in *Techniques and Instrumentation for Detection of Exoplanets VI*, Vol. 8864, International Society for Optics and Photonics, 88641N
- Pope, B., Cvetojovic, N., Cheetham, A., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 440, 125
- Por, E. H. 2017, in *Techniques and Instrumentation for Detection of Exoplanets VIII*, Vol. 10400, International Society for Optics and Photonics, 104000V
- Por, E. H., Haffert, S. Y., Radhakrishnan, V. M., et al. 2018, in *Proc. SPIE*, Vol. 10703, Adaptive Optics Systems VI
- Por, E. H. & Keller, C. U. 2016, in *Adaptive Optics Systems V*, Vol. 9909, International Society for Optics and Photonics, 990959

- Ruane, G., Ngo, H., Mawet, D., et al. 2019, *The Astronomical Journal*, 157, 118
- Sauvage, J.-F., Mugnier, L., Paul, B., & Villecroze, R. 2012, *Optics Letters*, 37, 4808
- Seldin, J. H., Paxman, R. G., Zarifis, V. G., Benson, L., & Stone, R. E. 2000, in *Imaging technology and telescopes*, Vol. 4091, International Society for Optics and Photonics, 48–63
- Snik, F. & Keller, C. U. 2013, in *Planets, Stars and Stellar Systems* (Springer), 175–221
- Snik, F., Otten, G., Kenworthy, M., et al. 2012, in *Modern Technologies in Space- and Ground-based Telescopes and Instrumentation II*, Vol. 8450, International Society for Optics and Photonics, 84500M
- Sparks, W. B. & Ford, H. C. 2002, *The Astrophysical Journal*, 578, 543
- Stam, D., Hovenier, J., & Waters, L. 2004, *Astronomy & Astrophysics*, 428, 663
- Traub, W. A. & Oppenheimer, B. R. 2010, *Exoplanets*, 111
- van Holstein, R. G., Snik, F., Girard, J. H., et al. 2017, in *Techniques and Instrumentation for Detection of Exoplanets VIII*, Vol. 10400, International Society for Optics and Photonics, 1040015
- Vigan, A., Gry, C., Salter, G., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 454, 129
- Vigan, A., N'Diaye, M., Dohlen, K., et al. 2018, in *Adaptive Optics Systems VI*, Vol. 10703, International Society for Optics and Photonics, 107035O
- Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, *Computing in Science & Engineering*, 13, 22
- Wilby, M. J., Keller, C. U., Snik, F., Korkiakoski, V., & Pietrow, A. G. 2017, *Astronomy & Astrophysics*, 597, A112

Appendix A: Phase retrieval examples

In section 2 we explained how the sign of even phase aberrations can be retrieved using an asymmetric pupil amplitude. In this appendix we will give a more complete set of examples of how pupil symmetries, and APPs designed for them, respond to even and odd aberrations.

In the top row of Figure A.2 a we visually present Equations (1) to (7) for an even aperture ($A = A_{\text{even}}$) and the odd coma aberration ($\theta = \theta_{\text{odd}}$). The real part of the pupil-plane electric field, $A \cos(\theta)$, is completely even and contains no sign information due to $\cos(-\theta) = \cos(\theta)$; the imaginary part, $iA \sin(\theta)$, is completely odd. Therefore, propagation to the focal-plane will result in a fully imaginary electric field ib (see Table 1) that has even and odd components. The bottom row of Figure A.2 a shows the same images but for a flipped sign of the phase ($\theta = -\theta_{\text{odd}}$), equivalent to conjugating the pupil-plane electric field. Thus the phase conjugation results in a flipped PSF, and due to the odd component of E_{foc} , it results in a morphological change that can be measured, which is clearly visible in Figure A.2 a.

Figure A.2 b shows similar figures, but for the even astigmatism aberration ($\theta = \theta_{\text{even}}$) and again, the top and bottom rows show opposite signs of the aberration. As expected, both real and imaginary parts of the pupil-plane electric field, $A \cos(\theta)$ and $iA \sin(\theta)$, are even. Again, the sign information is encoded in the imaginary pupil electric field term $iA \sin(\theta)$. This term propagates, in contrast to the odd phase aberration, to the real part of the focal-plane electric field a . Therefore, all sign information of the even phase aberration is encoded in the real focal-plane electric field. Because the aberration is even, this does not result in morphological changes in the total intensity measurement I_{foc} and thus sign information of the even aberration cannot be measured.

This well-known sign ambiguity (Gonsalves 1982; Paxman et al. 1992) is often broken by introducing a static, known diversity phase that results in a known real focal-plane electric field. This real electric field can interfere with the sign information carrying real electric field resulting from the even phase aberrations, which enables unique intensity changes for sign changes. Figure A.1 shows the case of defocus as a static phase diversity and the even astigmatism aberration used in Figure A.1. It shows a significant morphology change in both the real and imaginary focal-plane power $|a|^2$ and $|b|^2$ that can be used to determine sign changes.

In section 2 we explained how the Asymmetric Pupil Fourier Wavefront Sensor (Martinache 2013) breaks the sign ambiguity in a similar way using odd pupil amplitude. In Figure A.2 c the response of an asymmetric aperture to sign changes in odd phase aberrations is shown. It shows that, as expected, the ability of measuring odd phase modes is not affected. Figure A.2 d has an even phase aberration, again the top and bottom rows differ in the sign. In contrast to Figure A.2 b, the PSF for an asymmetric aperture does show a morphological change when the sign of the even aberration changes and therefore allows for the retrieval of the complete pupil-plane phase.

Now we will show that wavefront sensing with APPs will give similar results. In Figure A.3 a and b, the focal-plane intensity responses are shown to an odd and even pupil-plane phase aberration for an APP designed for a symmetric aperture. Similar to Figure A.2, the APP shows an intensity morphology change for a sign change in the odd phase aberration (Figure A.3 a), but for a sign change in the even phase aberration there is no

intensity morphology change (Figure A.3 b).

This changes for an APP designed for an asymmetric aperture, as shown in Figure A.3 c and d. Figure A.3 d shows the focal-plane intensity response to an odd pupil-plane phase aberration and Figure A.3 d the response to an even pupil phase aberration. Now, for both phase aberrations there is an intensity morphology change when the sign of the aberration changes.

Appendix B: Implications for spatial LDFC

Above, we discussed the principle behind FPWFS in the context of the coronagraphic PSFs of the APP, but the same principle applies to maintaining the contrast in the dark hole, such as electric field conjugation (Groff et al. 2015), speckle nulling (Bordé & Traub 2006) or spatial Linear Dark Field Control (LDFC; Miller et al. 2017). Spatial LDFC maintains the contrast in the dark hole by monitoring the intensity of the bright field, which can be shown to have a linear response to small phase aberrations. Suppose the focal-plane electric field E_{foc} consists of the nominal electric field in the bright field E_0 disturbed by the electric field of the aberration E_{ab} :

$$E_{\text{foc}} = E_0 + E_{\text{ab}}, \quad (\text{B.1})$$

$$I_{\text{foc}} = |E_{\text{foc}}|^2 \quad (\text{B.2})$$

$$= |E_0|^2 + |E_{\text{ab}}|^2 + 2\Re\{E_0 E_{\text{ab}}^*\}. \quad (\text{B.3})$$

Assuming that $|E_0|^2 \gg |E_{\text{ab}}|^2$, we can write the intensity change ΔI , compared to the reference image $I_0 = |E_0|^2$, due to the aberration as:

$$\Delta I = I_{\text{foc}} - I_0 \quad (\text{B.4})$$

$$= 2\Re\{E_0 E_{\text{ab}}^*\} \quad (\text{B.5})$$

Writing the electric fields as their real and imaginary components, $E_0 = a + ib$ and $E_{\text{ab}} = c + id$, shows that in order to have a response to the complete electric field of the aberration, the bright field of the PSF should have real and imaginary components:

$$\Delta I = 2(ac + bd). \quad (\text{B.6})$$

As discussed above, the real component a can be provided by either a known, even phase aberration such as defocus, or an amplitude asymmetry.

In Miller (2018) spatial LDFC was tested, both in simulation and in the lab, with a vAPP designed for an even aperture ($a = 0$). It was indeed observed that the LDFC loop was more stable when the vAPP image was defocussed ($a \neq 0$), compared to a focussed image ($a = 0$). Therefore, adding a pupil amplitude asymmetry, generating a non-zero a , would allow the LDFC loop to run with a focussed vAPP image, having a comparable stability as the defocussed one.

Appendix C: Derivatives objective function

The phase estimation is performed by minimizing the objective function \mathcal{L} defined in Equation 20. The convergence speed and accuracy of the minimization algorithm are improved when it is provided with the gradients of the parameters it is estimating. For the estimation, the phase is expanded on a truncated mode basis $\{\phi_i\}$ with α_i the estimated coefficients; thus the gradient $\partial \mathcal{L} / \partial \alpha_i$ needs to be derived. Besides the phase estimation, the algorithm can also estimate the photon number N_p , the background level N_b , the fractional degree of circular polarization v

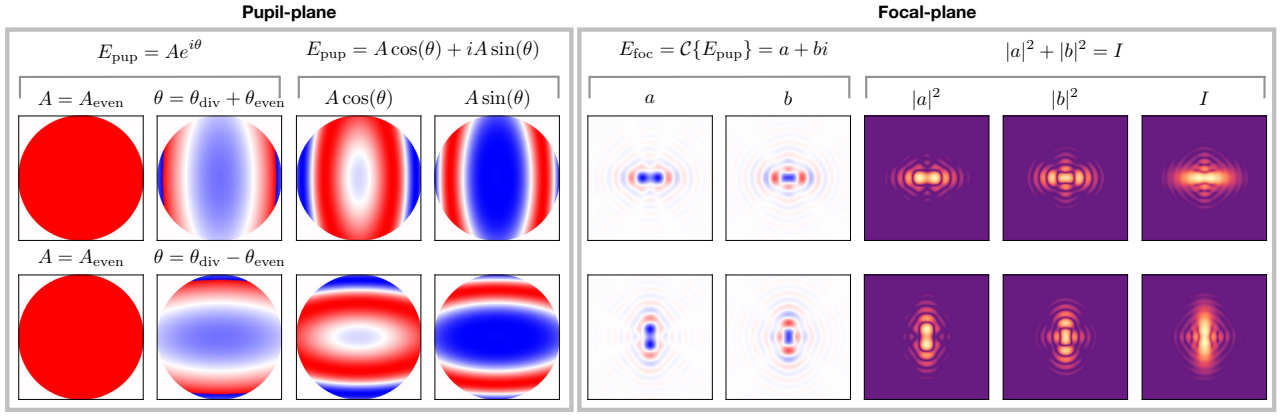


Fig. A.1. Focal- and pupil-plane quantities for a static diversity phase (defocus) with an even phase aberration (astigmatism) that has an alternating sign between the rows. The columns in the pupil-plane box show (from left to right) the amplitude, phase, real and imaginary electric field. In the focal-plane box, the columns show the real and imaginary electric fields, the power in the real and imaginary electric fields, and the total power.

and the amount of leakage L . Therefore, the gradients $\partial\mathcal{L}/\partial N_p$, $\partial\mathcal{L}/\partial N_b$, $\partial\mathcal{L}/\partial v$ and $\partial\mathcal{L}/\partial L$ also need to be derived.

The gradient of \mathcal{L} to X ($X = \{\alpha_i, N_p, N_b, v, L\}$) is given by:

$$\frac{\partial\mathcal{L}}{\partial X} = \sum_x \frac{1}{\sigma_n^2} [D - M] \frac{\partial M}{\partial X} + \frac{\partial\mathcal{R}(\alpha)}{\partial X}. \quad (\text{C.1})$$

with M the model of the system given by Equation 22. Here, the dependency of \mathcal{L} and M on (α, N_p, N_b, v, L) is omitted for readability. The term $\mathcal{R}(\alpha)/\partial X$ is only non-zero for $X = \alpha_i$.

We start with the gradient $\partial\mathcal{L}/\partial\alpha_i$; with the first step already shown in Equation C.1, the derivative of M is:

$$\frac{\partial M}{\partial\alpha_i} = N_p \sum_{j=1}^3 a_j(v, L) \frac{\partial I_{\text{foc},j}}{\partial\alpha_i} \quad (\text{C.2})$$

where the sum is over the two coronagraphic PSFs and the non-coronagraphic leakage PSF, and $I_{\text{foc},j}$ is given in Equation 26. The derivative $\partial I_{\text{foc},j}(\alpha)/\partial\alpha_i$ is:

$$\frac{\partial I_{\text{foc},j}(\alpha)}{\partial\alpha_i} = \frac{\partial E_{\text{foc},j}}{\partial\alpha_i} E_{\text{foc},j}^* + E_{\text{foc},j} \frac{\partial E_{\text{foc},j}^*}{\partial\alpha_i}, \quad (\text{C.3})$$

The partial derivatives to the focal-plane electric field E_{foc} are:

$$\begin{aligned} \frac{\partial E_{\text{foc},j}}{\partial\alpha_i} &= \frac{\partial}{\partial\alpha_i} C\{E_{\text{pup},j}(\alpha)\} \\ &= \frac{\partial}{\partial\alpha_i} C\{A e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\} \\ &= C\{A \frac{\partial}{\partial\alpha_i} e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\}, \text{ (Leibniz's integration rule)} \\ &= iC\{A \phi_i e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\} \end{aligned} \quad (\text{C.4})$$

The partial derivatives to the focal-plane electric field E_{foc}^* is then simply:

$$\frac{\partial E_{\text{foc},j}^*}{\partial\alpha_i} = -iC\{A \phi_i e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\}^* \quad (\text{C.5})$$

Combining these results, we find for $\partial I_{\text{foc},j}(\alpha)/\partial\alpha_i$:

$$\begin{aligned} \frac{\partial I_{\text{foc},j}(\alpha)}{\partial\alpha_i} &= iC\{A \phi_i e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\} C\{A e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\}^* \\ &\quad - iC\{A e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\} C\{A \phi_i e^{i(\theta_j + \sum_k \alpha_k \phi_k)}\}^* \end{aligned} \quad (\text{C.6})$$

The regularization term $\mathcal{R}(\alpha)$ that we adopted is fairly simple:

$$\mathcal{R}(\alpha) = \frac{1}{2} \sum_{k=1}^N \frac{\alpha_k^2}{k^\gamma}, \quad (\text{C.7})$$

with γ a power that matches the known or assumed power-law distribution of the aberrations in the system. The derivative with respect to α_i is given by:

$$\frac{\partial\mathcal{R}(\alpha)}{\partial\alpha_i} = \frac{\alpha_i}{i^\gamma} \quad (\text{C.8})$$

Combining the results in Equation C.1, Equation C.2, Equation C.6 and Equation C.8 gives the final expression for $\partial\mathcal{L}/\partial\alpha_i$.

Next, the derivatives of M to N_p and N_b are:

$$\frac{\partial M}{\partial N_p} = \sum_{j=1}^3 a_j(v, L) I_{\text{foc},j}(\alpha), \quad (\text{C.9})$$

$$\frac{\partial M}{\partial N_b} = 1 \quad (\text{C.10})$$

The derivative of M to the leakage L is:

$$\frac{\partial M}{\partial L} = N_p \sum_{j=1}^3 \frac{\partial a_j(v, L)}{\partial L} I_{\text{foc},j}, \quad (\text{C.11})$$

with the derivatives $\partial a_j(v, L)/\partial L$ given by:

$$\frac{\partial a_1}{\partial L} = -\frac{1+v}{2}, \quad (\text{C.12})$$

$$\frac{\partial a_2}{\partial L} = \frac{v-1}{2}, \quad (\text{C.13})$$

$$\frac{\partial a_3}{\partial L} = 1. \quad (\text{C.14})$$

Finally, the derivative of M to v is:

$$\frac{\partial M}{\partial v} = N_p \sum_{j=1}^3 \frac{\partial a_j(v, L)}{\partial v} I_{\text{foc},j}, \quad (\text{C.15})$$

with the derivatives $\partial a_j(v, L)/\partial v$ given by:

$$\frac{\partial a_1}{\partial v} = \frac{1-L}{2} \quad (\text{C.16})$$

$$\frac{\partial a_2}{\partial v} = \frac{L-1}{2} \quad (\text{C.17})$$

$$\frac{\partial a_3}{\partial v} = 0 \quad (\text{C.18})$$

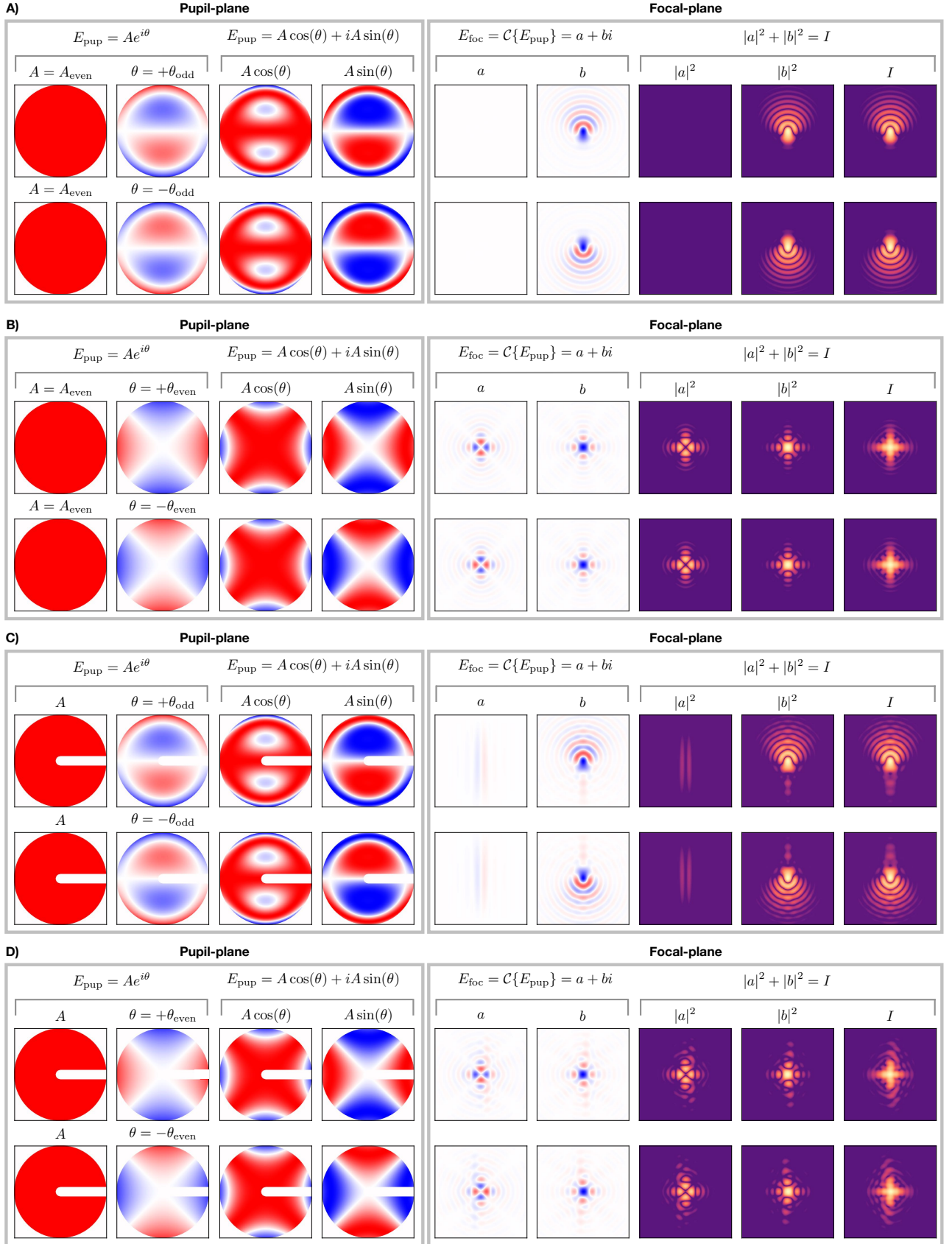


Fig. A.2. Focal- and pupil-plane quantities for different combinations of pupil symmetries and phase aberrations with alternating sign. (a) Symmetric aperture with odd aberration coma. (b) Symmetric aperture with even aberration astigmatism. (c) Asymmetric aperture with odd aberration coma. (d) Asymmetric aperture with even aberration astigmatism. The columns in the pupil-plane box show (from left to right) the amplitude, phase, real and imaginary electric field. In the focal-plane box, the columns show the real and imaginary electric fields, the power in the real and imaginary electric fields, and the total power.

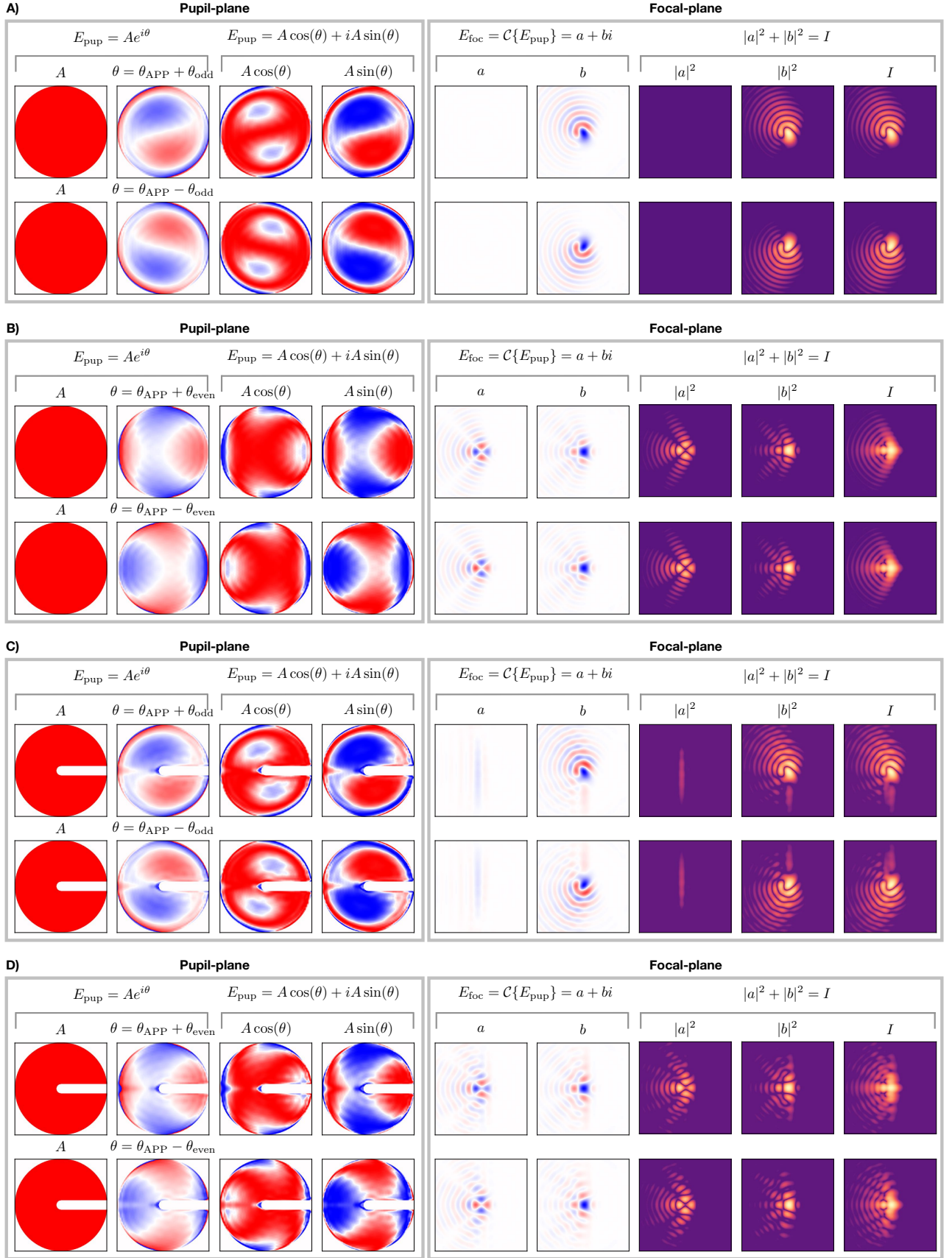


Fig. A.3. Focal- and pupil-plane quantities for APPs designed for different pupil symmetries subjected to phase aberrations with alternating sign. (a) Symmetric aperture with odd aberration coma. (b) Symmetric aperture with even aberration astigmatism. (c) Asymmetric aperture with odd aberration coma. (d) Asymmetric aperture with even aberration astigmatism. The columns in the pupil-plane box show (from left to right) the amplitude, phase, real and imaginary electric field. In the focal-plane box, the columns show the real and imaginary electric fields, the power in the real and imaginary electric fields, and the total power.